ADAPTIVE CODED MODULATION CLASSIFICATION AND SPECTRUM SENSING FOR COGNTIVE RADIO SYSTEMS

A. O. A. AL-JUBOORI

Ph.D

UNIVERSITY OF BRADFORD

2018

ADAPTIVE CODED MODULATION CLASSIFICATION AND SPECTRUM SENSING FOR COGNITIVE RADIO SYSTEMS

Adaptive Coded Modulation Techniques for Cognitive Radio Using Kalman Filter and Interacting Multiple Model Methods

Ahmed O. Abdul Salam Al-Juboori

BSc, MSc

Submitted for the Degree of

Doctor of Philosophy

Faculty of Engineering and Informatics

University of Bradford

2018

ABSTRACT

ADAPTIVE CODED MODULATION, CLASSIFICATION AND SPECTRUM SENSING FOR COGNITIVE RADIO SYSTEMS

Adaptive Coded Modulation Techniques for Cognitive Radio Using Kalman Filter and Interacting Multiple Model Methods

Keywords

Adaptive Coded Modulation, Automatic Modulation Classification, Spectrum Sensing, Cognitive Radio, Kalman Filter, Interacting Multiple Model

The current and future trends of modern wireless communication systems place heavy demands on fast data transmissions in order to satisfy end users' requirements anytime, anywhere. Such demands are obvious in recent applications such as smart phones, long term evolution (LTE), 4 & 5 Generations (4G & 5G), and worldwide interoperability for microwave access (WiMAX) platforms, where robust coding and modulations are essential especially in streaming on-line video material, social media and gaming. This eventually resulted in extreme exhaustion imposed on the frequency spectrum as a rare natural resource due to stagnation in current spectrum management policies. Since its advent in the late 1990s, cognitive radio (CR) has been conceived as an enabling technology aiming at the efficient utilisation of frequency spectrum that can lead to potential direct spectrum access (DSA) management. This is mainly attributed to its internal capabilities inherited from the concept of software defined radio (SDR) to sniff its surroundings, learn and adapt its operational parameters accordingly. CR systems (CRs) may commonly comprise one or all of the following core engines that characterise their architectures; namely, adaptive coded modulation (ACM), automatic modulation classification (AMC) and spectrum sensing (SS).

Motivated by the above challenges, this programme of research is primarily aimed at the design and development of new paradigms to help improve the adaptability of CRs and thereby achieve the desirable signal processing tasks at the physical layer of the above core engines. Approximate modelling of Rayleigh and finite state Markov channels (FSMC) with a new concept borrowed from econometric studies have been approached. Then insightful channel estimation by using Kalman filter (KF) augmented with interacting multiple model (IMM) has been examined for the purpose of robust adaptability, which is applied for the first time in wireless communication systems. Such new IMM-KF combination has been facilitated in the feedback channel between wireless transmitter and receiver to adjust the transmitted power, by using a water-filling (WF) technique, and constellation pattern and rate in the ACM algorithm. The AMC has also benefited from such IMM-KF integration to boost the performance against conventional parametric estimation methods such as maximum likelihood estimate (MLE) for channel interrogation and the estimated parameters of both inserted into the ML classification algorithm. Expectation-maximisation (EM) has been applied to examine unknown transmitted modulation sequences and channel parameters in tandem. Finally, the non-parametric multitaper method (MTM) has been thoroughly examined for spectrum estimation (SE) and SS, by relying on Neyman-Pearson (NP) detection principle for hypothesis test, to allow licensed primary users (PUs) to coexist with opportunistic unlicensed secondary users (SUs) in the same frequency bands of interest without harmful effects. The performance of the above newly suggested paradigms have been simulated and assessed under various transmission settings and revealed substantial improvements.

Acknowledgment

I'm indebted to many people during the course of pursuing my research study. First and foremost, I would like to show my appreciation and gratitude for my supervisors Prof. Ray E. Sheriff and Prof. Saleh R. Al-Araji for introducing me to this vital subject and for their continuous guidance, support, and advice. Without their profound contributions, understanding, and being patient with me, the accomplishment of this study would not be easily realised.

I also would to like to extend my thanks to Prof. Yim F. Hu for taking over and supervised the last year of my thesis write-up when Prof. Ray Sheriff transited to another academic institution. Prof. Yim F. Hu dedicated support is greatly appreciated. Thanks are due to Dr. Kahtan Mezher and Dr. Qassim Nasir as well for their valuable participation in articulating several scholar publications.

The kind assistance and support of all faculty and staff members and post graduate students of the Faculty of Engineering and Informatics at the University of Bradford is generously acknowledged.

This context should not end without mentioning the generous compassion, understanding, and tolerance received from my wife Huda, and children Osama, Omar, and Sara. Their dearly presence in my life goes beyond any description.

My last words of appreciation are endlessly to my mother and father for their unconditional embracement, prayers and encouragement towards achieving my higher studies. I owe them so much in my entire life.

Table of Contents

Acronyms

4G	4th Generation
5G	5th Generation
3GPP	3rd Generation Partnership Project
xG	NeXt Generation
AM	Amplitude Modulation
ACF	Autocorrelation Function
ACK	Acknowledgment (positive)
ACM	Adaptive Coded Modulation
ADC	Analog to Digital Converter
AFRL	Air Force Research Laboratory
ALRT	Average LRT
AMC	Automatic Modulation Classification
ANN	Artificial Neural Network
APK	Amplitude-Phase Keying
AQAM	Adaptive QAM
AR	Autoregressive
ARMA	Autoregressive MA
ARQ	Automatic Repeat Request
ASE	Adaptive Spectrum Efficiency
ASHA	Adaptive Spectrum Hyperspace Access
ASIC	Application Specific Integrated Circuit
AT	Adaptive threshold
ATV	Analogue Television
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BF	Beamforming
BHT	Binary Hypothesis Test
BICM	Bit Interleaved Coded Modulation
BICM-ID	BICM with Iterative Decoding
BLUE	Best Linear Unbiased Estimator
BPF	Bandpass Filter
BPSK	Binary PSK
BS	Base Station
BSC	Binary Symmetric Channel
BSS	Blind Source Separation
BWA	Broadband Wireless Access
CC	Cognitive Cycle
CDF	Cumulative Distribution Function
CDR	Constant Detection Rate
CFAR	Constant False Alarm Rate
CLT	Central Limit Theory

CM	Coded Modulation
CPE	Customer Premises Equipment
CR	Cognitive Radio
CRLB	Cramer-Rao Lower Bound
CRN	Cognitive Radio Network
CSD	Cyclostationary Detector
CSI	Channel State Information
CSIR	CSI-Receiver
CSIT	CSI-Transmitter
DA	Data Aided
DAC	Digital to Analogue Converter
DARPA	Defence Advanced Research Projects Agency
DDCE	Decision Directed Channel Estimation
DoF	Degree of Freedom
DPSS	Discrete Prolate Spheroidal Sequence
DRVP	Discrete-Rate Variable-Power
DSA	Dynamic Spectrum Access
DSAN	DSA Network
DSM	Dynamic Spectrum Management
DSMC	Discrete-States Markov Chain
DSP	Digital Signal Processing (or Processors)
DTV	Digital Television
DySPAN	Dynamic Spectrum Access Networks
E2R	End to End Re-Configurability
ED	Energy Detector
EGC	Equal-Gain Combining
EKF	Extended KF
EM	Expectation-Maximisation
ET	Eigenmode Transmission
EU	European Union
EVD	Eigenvalue Decomposition
FB	Feature Based
FBk	Filter Bank
FBkSE	Filter Banks Spectrum Estimator
FCC	Federal Communications Commission
FEC	Forward Error Correction
FET	Frobenius ET
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FPGA	Field Programming Gate Array
FSMC	Finite State Markov Chain
GA	Genetic Algorithm
GLRT	General LRT
GMD	Geometric Mean Decomposition
GPP	General Purpose Processor
HARQ	Hybrid ARQ
HDL	Hardware Description Language
HLRT	Hybrid LRT

HMI	Human Machine Interface
HMM	Hidden Markov Model
HOS	High Order Statistics
ICA	Independent Component Analysis
ICNIA	Integrated Communications, Navigation, Identification Avionics
IEEE	Institute of Electrical and Electronics Engineers
IF	Intermediate Frequency
IFFT	Invers Fast Fourier Transform
IMM	Interacting Multiple Model
ISI	Inter-symbol Interference
ISM	Industrial, Scientific and Medical
ITU	International Telecommunications Union
JADE	Joint Approximate Diagonalization of Eigenmatrices
JCEDD	Joint Channel Estimation and Data Detection
JMLS	Jump Markov Linear System
KF	Kalman Filter
KLT	Karhunen-Loeve Transform
KNN	K-Nearest Neighbour
KQML	Knowledge Queue Manipulative Language
K-S	Kolmogorov-Smirnov
LB	Likelihood Based
LCR	Level Crossing Rate
LF	Likelihood Function
LLF	Log-Likelihood Function
LLR	Log-Likelihood Ratio
LMMSE	Linear MMSE
LMS	Least Mean Square
LOS	Line-of-Sight
LPF	Low-Pass Filter
LS	Least-Square
LSE	LS Estimator
LTE	Long Term Evolution
LRT	Likelihood Ratio Test
LSE	LS Estimator
LUT	Look-Up Table
MA	Moving Average
MAC	Medium Access Control
MAN	Metropolitan Area Network
MAP	Maximum A-Posteriori
MAPE	MAP Estimator
MED	Minimum Euclidean Distance
MF	Matched Filter
MHT	Multi-Hypotheses Test
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
ML	Maximum Likelihood
MLE	ML Estimate
MMS	Minimum Mean Square

MMSE	MMS Error
MoM	Method of Moments
MQAM	M-ary Quadrature Amplitude Modulation
MQPSK	M-ary Quadrature Phase Shift Keying
MRC	Maximum-Ratio Combining
MRT	Maximum-Ratio Transmitting
MTM	Multitaper Method
MTSE	Multitaper Spectrum Estimator
MU	Multiuser
MVUE	Minimum Variance Unbiased Estimator
NACK	Negative Acknowledgment
NDA	Non-Data Aided
NLOS	Non-Line-Of-Sight
NP	Neyman-Pearson
Ofcom	Office of Communication
OFD	Optimum Frame Duration
OFDM	Orthogonal Frequency Division Multiplexing
OOB	Out-Of-Band
OSA	Opportunistic Spectrum Access
OSI	Open System Interface
PACE	Pilot-Assisted (or Aided) Channel Estimation
PCA	Principal Component Analysis
PCM	Pulse-Code Modulation
PDF	Probability Density Function
PER	Pairwise Error Rate
PHY	Physical Layer
PLC	Power Line Communication
PSAM	Pilot Symbol Assisted Modulation
PSD	Power Spectral Density
PSE	Periodogram Spectral Estimator
PSK	Phase Shift Keying
PU	Primary User
QAM	Quadrature Amplitude Modulation
Q(H)LRT	Quasi(H)LRT
QLD	Quadrature Left Decomposition
QoS	Quality of Service
QPSK	Quadrature PSK
QRD	Quadrature Right Decomposition
QSTBC	Quasi-STBC
RBLS	Rao-Blackwell-Lehmann-Scheffe
RF	Radio Frequency
RKRL	Radio Knowledge Representation Language
RLS	Recursive LS
RMSE	Root-Mean Square Error
ROC	Receiver Operating Characteristic
SC	Selection Combining
SCC	Standard Coordination Committee
SDR	Software Defined Radio

SER	Symbol Error Rate
SIC	Successive Interference Cancellation
SIMO	Single-Input Multiple-Output
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SM	Spatial Multiplexing
SNR	Signal-to-Noise Ratio
SRSV	Square-Root Singular Value
SS	Spectrum Sensing
SSC	Switch-and-Stay Combining
STBC	Space-Time Block Code
STMF	Space-Time MF
STTC	Space-Time Trellis Code
SU	Secondary User
SVD	SV Decomposition
SVM	Support Vector Machine
TCM	Trellis Coded Modulation
TD-LTE	Time-Division LTE
TTCM	Turbo TCM
TuCM	Turbo Coded Modulation
TV	Television
TVWS	Television White Spaces
UHF	Ultra-High Frequency
UK	United Kingdom
UMP	Uniformly Most Powerful
USA	United States of America
USRP	Universal Software Radio Peripheral
UWB	Ultra-Wide Band
VA	Viterbi Algorithm
V-BLAST	Vertical-Bell Labs Layered Space-Time
VHDL	Verilog HDL
VHF	Very-High Frequency
VQ	Vector Quantisation
VRVP	Variable-Rate and Variable-Power
WF	Water-Filling
WG	Working Group
WiFi	Wireless Fidelity
WiMAX	Worldwide Microwave Access
WLAN	Wireless Local Area Network
WOSA	Weighted Overlapped Segment Averaging
WPAN	Wireless Personal Area Network
WRAN	Wireless Rural Area Network
WSS	Wide Sense Stationary
ZMCSCG	Zero-Mean Circularly Symmetrical Complex Gaussian

List of Figures

List of Tables

Mathematical Notations

$\boldsymbol{x}$, $\boldsymbol{y}$	boldface lowercase letters denote vectors
$\boldsymbol{H}$, $\boldsymbol{Y}$	boldface uppercase letters denote matrices
h, a, X	non-boldface upper and lower case letters denote scalars
$\Re(.)$, $\Im(.)$	real and imaginary parts of a process, also denoted by $\mathrm{Re}(.)$, $\mathrm{Im}(.)$
$min(.)$	minimisation function
$max(.)$	maximisation function
$sin(.)$	sine function
$ln(.)$	natural logarithm function
$log_x(.)$	logarithm function of base x
$exp(.)$	exponential function, also denoted by $e^{(.)}$
FFT{.}	fast Fourier transform
IFFT{.}	inverse fast Fourier transform
$\mathcal{CN}(.,.)$	circular Gaussian complex process with arbitrary mean and variance
Q(.)	right-tail CDF given by $\mathrm{Q}(.) = \frac{1}{\sqrt{2\pi}}\int_{.}^{\infty} e^{-x^2/2}dx$
$erfc(.)$	complementary error function defined as $\mathrm{Q}(.) = \frac{1}{2}erfc(\frac{.}{\sqrt{2}})$
$erf(.)$	error function equals to 1- $erfc(.)$
$F(.)$	CDF of a RV
$p(.)$	PDF of a RV
Pr(.)	probability of a RV
$\frac{\partial y}{\partial x}$	first derivative
$\frac{\partial^2 y}{\partial x^2}$	second derivative
$\sum.$	summation, also denoted by $\mathrm{sum}(.)$
$\prod.$	multiplication
$\mathcal{O}(.)$	mathematical operations (summations, subtraction, multiplication, division)
$j = \sqrt{-1}$	imaginary operator
$(.)^{\mathrm{T}}$	transpose operator
$(.)^{*}$	complex conjugate
$(.)^{\mathrm{H}}$	Hermitian operator
$\|.\|$	vector norm

$\|\|.\|\|_F$	Frobenius norm
$\mathbb{E}(.)$	expectation operator
$\mathrm{Tr}(.)$	trace operator
$\det(.)$	determinant operator
$\mathrm{vec}(.)$	vector stacking of a vector process
$\mathrm{diag}[.,.,.]$	diagonal entries of a matrix
$\rightarrow$	mapping from one space to another
$\forall$	elements belong to, or $\in$, a certain space, sequence or range $[.,.]$, $\{.\}$, $(.,.)$

Chapter 1

INTRODUCTION

1.1 Rationale

The current trend of modern wireless communication systems is to embed more intelligent features to adaptively enhance transmission performance and qualities, for example; minimising power and bandwidth and increasing reliability. This has generated a drive for developed systems to have flexibility in transmission features such as efficient spectrum allocation, output power, coding and modulation schemes. Despite improving performance, these technologies entail the radio frequency (RF) spectrum to be dynamically allocated and consequently its controllability becomes challenging as more devices and services are added. On a slower tandem, however, existing regulation and governing policies are expected to continue over the predictable future. To this extent, the ability to adaptively determine spectrum spaces, coding and modulation types allows to achieve cognition cycles at affordable overheads over conventional radio networks.

Cognitive radio (CR) as an appealing multidisciplinary concept in contemporary wireless networks has extensively grown in the last two decades. The cognition feature in modern radio systems enhances the overall performance by adapting to changing transmission and reception environments. CRs promote the awareness against wider aspects of external variations by having better flexibility to cope with the internal hardware. The new paradigm of CR networks (CRNs) embracing the envisioned software-defined-radio (SDR) and targeting smarter and efficient wireless spaces has thus been evoked.

This chapter provides the fundamental background and aspirations that inspired the motivation and impact of this study programme. This is illustrated by having an insight into the current RF spectrum utilisation and its need to be viably more efficient by employing feasible adaptation techniques such as CR systems. This chapter also outlines the main knowledge contributions and chapters' organisation of this study programme.

1.2 Motivation and Background

Today's world is on the verge of explosive demand for a broadband spectrum to support wireless streamline applications and online services demanded by diverse users and smart devices irrespective of their geographical location and at any time. This has imposed a heavy demand for reliable and efficient utilisation of the currently available spectrum and to explore new bands in the upper RF range for possible exploitation. Recent studies, however, have shown that the currently permissible RF spectrum is intensively underutilised in many respects. One of the prime reasons of such underutilisation is attributed to the long-term static lease of certain frequency bands assigned for licensed primary users (PUs) over large topographical areas [1, 2]. Such stagnation in the current spectrum access and management policies has resulted in a negative impact on spectrum in the urban and suburban areas due to the following factors over long periods of time:

- Some spectrum bands are largely vacant;
- Some spectrum bands are sparsely occupied;
- The few remaining spectrum bands are severely utilised.

Recently, there have been several worldwide site measurement campaigns endeavouring to realistically quantify spectrum occupancy levels and thereby help decision makers to maintain appropriate strategic planning for such scarce natural resource. For example, the Federal Communications Commission (FCC) has indicated that variations in the utilisation of assigned spectrum across time and space range from 15% to 85% in the United States of America (USA) [2]. Severely deteriorated spectrum underutilisation figures as low as 4.54% in the operative bands in other countries have also been reported [3].

Such inferiorly managed spectrum utilisation has led to the evolvement of a new concept traditionally known as spectrum holes or white spaces [1]. This would mandate for an efficient spectrum management and notable improvement over prevalent radio systems performance. This can possibly be achieved by allowing unlicensed secondary users (SUs) to approach the white spaces unused by PUs at the right location and time.

This accordingly contributed to the emergence of CR technology to improve the spectrum availability and accessibility.

Television (TV) bands, at very-high frequency (VHF) and ultra-high frequency (UHF), respectively, for their desirable transmission characteristics have attracted tremendous attention to begin with for serving new spectrum sharing policies. Such bands of interest, commonly dubbed TV white spaces (TVWS) [4-7], are considered decent contenders for alternative solutions based on CR technology operating on a license-exempt basis in TVWS spectrum. Their commercial appeal is well evolving, particularly for fixed line operators that have a significant fibre and copper infrastructure, in addition to potential new players like Google and Microsoft. The TVWS are well anticipated to deliver viably high and scalable solutions compared to other conventional options such as cellular and/or wireless fidelity (WiFi) technologies comprised in most IEEE 802.11 family of standards.

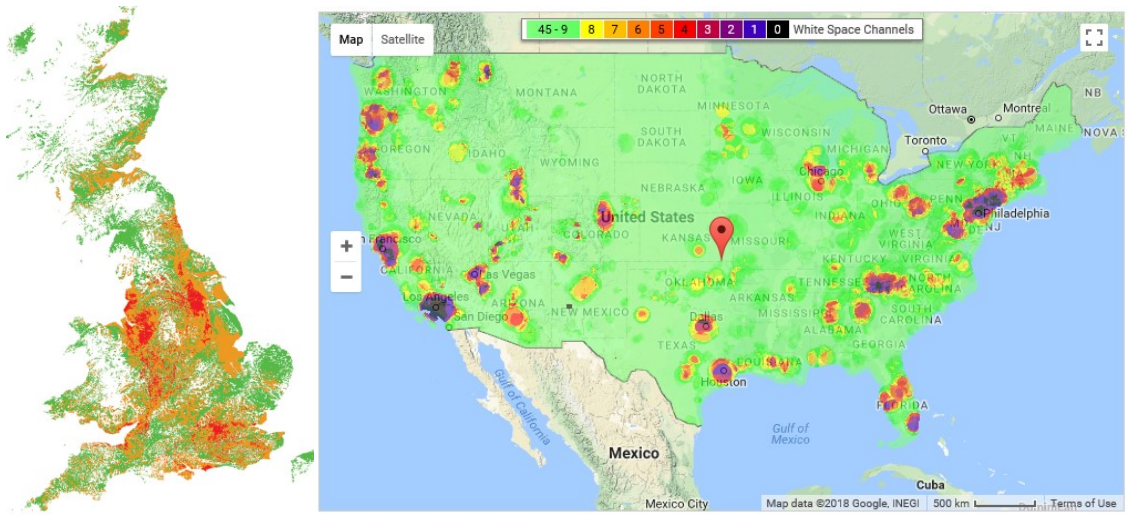


Fig. 1.1 TVWS distribution across the UK (left) and USA (right) red (purple): little; white (green): much (Courtesy [4] and Google).

The TVWS prodigy occurs as a side product of the digital switchover given that the transition from analogue TV (ATV) to digital TV (DTV) has already taken place in most countries and in progress in others. After completing the digital switchover, a large portion of RF spectrum within the existing TVWS will become available for sharing using innovative strategies such as dynamic spectrum access or allocation (DSA) [2]. A map of the United Kingdom (UK) showing TVWS anticipated after the

digital switchover is depicted above in Fig. 1.1 (left) [4]. It can be seen that massive TVWS are available in remote areas like Wales and Scotland. Likewise, the free interactive map of TVWS in the USA can be accessed on the Google site at any time, and compatible results were shown in earlier study [7]. Further guidelines on achievable TVWS capacities and technical challenges can be found in [6, 7].

In an effort to take full advantage of the abundant vacant spectrum available in the TVWS bands, the FCC in the USA from 2004 to 2008 has developed and then approved the IEEE 802.22 to grant permission for fixed broadband systems to access these bands for industrial and commercial purposes [6-8]. Such free SUs access must be well distinguished from both unlicensed and currently used industrial, scientific, and medical (ISM) bands and is subject to the condition that no harmful or interference effects should be caused to PUs occupying the same bands whatsoever. It is worthy to note that TVWS bands are more desirable and convenient for many wireless transmission services due to their superior propagation characteristics. These bands reside below 1 GHz frequency; scattering and obstruction are less harmful than at higher frequencies, allowing non-line-of-sight (NLOS) coverage. Also the path loss of TVWS bands is more advantageous over ISM bands (2.4 and 5.7 GHz) due only to operating frequency.

Office of Communication (Ofcom), the UK's independent spectrum regulator, has been at the forefront of promoting secondary access to TVWS in Europe. Ofcom hence followed the lead of the FCC and issued its first statement in 2007 and followed by a consultation in 2009. It then released the operation permission at the end of 2012 by legalizing the use of interleaved TVWS spectrum by license-exempt or exempt-exempt devices [4, 5]. Ofcom, in a later step and in an effort to establish interference avoidance to PUs, has evaluated three mechanisms for identification of vacant TVWS bands, namely; spectrum sensing (SS), geolocation database and beacons. The geolocation database and beacons' mechanisms maintain spatiotemporal separation between PUs and SUs by using cooperative information on time, location, signal power and pilot guidance; the SS on the other hand allows for safe coexistence and can be conducted locally and blindly. The SS and other adaptation mechanisms are of prime interest in this research and hence will be briefly outlined below. Additional details can be found in the subsequent chapters.

1) Spectrum Sensing (SS)

The SS mechanism is considered as a fundamental feature in the building architectures of CRs and other cognitive devices [5-8]. It involves continuously monitoring the RF spectrum and processing the incoming information to seek for possible vacant channels not admitted by licensed PUs in order to allow opportunistic SUs the possibility to occupy such empty channels. The requirements of this mechanism are mainly determined by four central parameters: sensing receiver sensitivity, channel detection time, probability of detection, and probability of false alarm. The sensing receiver sensitivity is -116 dBm for DTV, -94 dBm for ATV, while for wireless microphones the sensitivity is -107 dBm. The channel detection time for all signal types is 2 Seconds. The minimum probability of detection is 0.9, while the maximum probability of false alarm is 0.1 for all signal types [8].

2) Adaptive Coded Modulation (ACM)

In the literature, this mechanism is also broadly known as adaptive modulation and coding, but the term ACM is used to avoid confusion with the third mechanism described below. Once the best available channel has been selected by a CR node, the next functional challenge is to make the network protocols adaptive to the available spectrum [2]. The ACM mechanism is aimed at the adaptive resource allocation to improve the overall performance of a CR communication system. A CR equipped with the ACM feature has the ability to alter its internal parameters based on the outputs of active transmission monitoring such as radio spectrum, licensed PUs' activities and traffic and fading channel variations. The IEEE 802.22 has supported 12 combinations of 3 modulations; namely, quadrature phase shift keying (QPSK), 16-quadrature amplitude modulation (QAM) and 64-QAM, and 4 coding rates (1/2, 2/3, 3/4, 5/6) for data communications. A CR can flexibly choose among the above combinations to achieve various tradeoffs of data rate and robustness, depending on channel state information (CSI) and interference conditions [5].

3) Automatic Modulation Classification (AMC)

The AMC, or occasionally called automatic modulation recognition, despite its military roots, is a vital mechanism to support the online selection of suitable demodulation process in the CR receiver [9, 10]. Whether having a little or no prior

knowledge of the received signal characteristics, a CR must be able to recognise the RF spectrum accurately. The symbol rate, carrier frequency, coding and modulation types can be among other important parameters that need to be accurately identified to boost the CR awareness against its surround. An important aspect of interference avoidance can thus be achieved by having coexisting cognitive SUs constantly sensing the transmission pattern of PUs in order to be dynamically controlled based on the AMC protocol. Fusing the AMC mechanism with SDR platforms is projected to form CR engines necessary for efficient SS schemes. Such modernised single chips are capable of sensing and detecting all signal forms of PUs and eventually lead to fruitful operations of DSA schemes without compromising the overall CRNs' performance as per prevalent spectrum access rules [9].

1.3 Aims and Objectives

As shown earlier in this context, the main strategic objectives of CR principle retained wide doors open for further investments in exploring new application fields and genuine thoughts. Recent advancements strived considerably to attain higher levels of quality of service (QoS) imposed by demanding users adopting contemporary mass communication tools as a daily habit for their lives. Such trends are exponentially rising and ultimately revolutionizing the modern life without leaving any turning backs. The CR research endeavours are thus obliged to keep pace with these technical challenges and motivations and strive to surpass expectations substantially.

Acknowledging the above given facts and challenges, this study aims to meet manifold objectives stated below while keeping in mind the increasing engagement of CR systems in a wider aspects of modern applications. This study programme is hence mainly concerned with the design and development of new paradigms that have a direct impact on the efficient utilisation of frequency spectrum by interrogating baseband signals at the physical layer (PHY) of CR systems. The main goals as such are to advocate viable spectrum sustainability as much as possible while addressing the accessible technologies at present.

To achieve the main goals of this research programme, certain objectives have been set:

- **First objective**: is to elaborate on the best approximate modelling of wireless fading channels, which constitutes the prime transmission medium of CR systems. The result of such channel modelling highly helps for better understanding of the requirements involved in the design and development of the three main mechanisms; namely, the SS, ACM, and the AMC.

- **Second objective**: is to come up with a proposal for robust channel fading estimation and thereby supply with the CSI attributes necessary for a transmitter to adjust its parameters accordingly. Such CSI estimate adaptation is crucial to all CR functionalities and helps achieve sustainable spectrum utilisation and avoids any compromise to PUs presence.

- **Third objective**: is to facilitate the adaptive CSI estimation feedback to improve the ACM operation by adjusting the necessary transmission parameters such as power, modulation and coding patterns. This, of course, can only occur when there is cooperation between both sides of a CR system, i.e., transmitter and receiver.

- **Fourth objective**: is to investigate AMC for introducing a blind measure of signal detection and classification by using statistical methods and without having any prior knowledge on the received digital sequences. This stimulating objective can be furtherly emphasised by exploiting the adaptively estimated CSI parameters to significantly improve such task.

- **Fifth and final objective**: is aimed at the multifaceted modelling of SS mechanism and analysis for the purpose of having a good trade-off between spectrum bias and leakage within the designated bands of interest for efficient detection of PUs' activities.

1.4 Research Importance and Contributions

The analytical procedures and outcomes of this study programme are well perceived to have a constructive contribution to the rapidly expanding world of CR systems. The scarce natural resource of frequency spectrum is on an unprecedented brink of severe fatigue and the entire world strives to explore new paradigms and alternative routes to

make the best of it. Innovative ideas to push the spectrum boundaries to higher frequency ranges unexplored before can also be noticed nowadays and anticipated to have more attention in the near future. Users' mobility and unpredicted emerging applications are almost eminent on a daily basis at the expense of frequency spectrum, which unfortunately remains limited. Therefore, this study and other efforts elsewhere endeavour to invest on the current spectrum situation by supporting regulators to find means for efficient spectrum utilisation and try to keep it green without negative consequences on the natural resources. The outcome of the study has helped for better understanding of the current constraints and affordable remedial options within the perimeter of reachable technologies available nowadays.

So much as that briefly described the origins of relevance; this study has major contributions that can be summarised as given below:

- Examined and analysed the conventional fading channels using Rayleigh statistics, autoregressive (AR) and finite state Markov channel (FSMC) modelling and for the first time employed a new governing structure borrowed from the econometric studies.

- Adopted Kalman filter (KF) and combined with interacting multiple model (IMM) as robust CSI estimation arrangement, which is newly introduced to this application as it is predominately used before in the air and vehicular traffic management areas.

- Exploited the IMM-KF integrated model as an efficient feedback originated from a CR receiver to deliver important information on the CSI status for transmitter to modify its ACM attributes accordingly. This is attained by suitably adjusting the signal power, modulation and coding, which is an innovative idea not examined earlier.

- Designed an AMC scheme using likelihood-based (LB) approach and the newly introduced IMM-KF algorithm for resilient expectation-maximisation (EM) procedure to effectively produce improved correct-classification rates.

- Developed a new adaptive SS scheme based on multitaper method (MTM) for robust spectrum control and used the IMM-KF to adjust for necessary sequence

length and decision threshold necessary for efficient PUs' presence detection to either occupy or vacate designated channels by SUs accordingly.

1.5 Thesis Organisation

This thesis is delivered in eight chapters. This chapter has supported a short review for the background of CRs and particularly stressed their role for efficient spectrum access and utilisation using the newly authorised standards and policies in this regard. The significant challenges, objectives and importance of this study programme are also briefly presented along with the summary of achievements in this chapter. Matlab simulations are administered to evaluate the performance of analytical algorithms and procedures in all chapters. The contents of remaining chapters are summarised below:

- **Chapter Two:** Unveils the CR aspects in terms of operational boundaries and reviews the state of the art in this vibrant domain. Other standards warranting CR functions in their architectures are also identified in this chapter.

- **Chapter Three:** Exposes the foundations of digital signalling and transmission techniques in terms of M-ary modulations, multiple antennas and carriers such as multiple-input multiple-output (MIMO), orthogonal-frequency division multiplexing (OFDM), Alamouti and space-time block codes (STBC) techniques. Also furnishes the traditional combining and singular-value decomposition (SVD) techniques necessary for the subsequent chapter developments. Finally, it covers the essential modelling requirements of wireless fading channels using the common Rayleigh distribution and Gilbert-Elliot scheme for Doppler correlation effect. The 1^{st}-order of both AR and FSMC models are examined to facilitate the Yule-Walker equations and a new scheme devised by Tauchen in the econometric studies is acquired to control the transition convergence among FSMC neighbour and far states.

- **Chapter Four:** The state space modelling necessary for KF is developed along with other conventional methods like least square (LS) and minimum mean square (MMS) for CSI estimate. The adaptation of KF is empowered by employing the vibrant IMM algorithm, which corresponds to the FSMC transition among various KFs' states. Complete analysis of the newly introduced

combination of IMM-KF scheme is provided in terms of banks of several KFs that each imitates different noise and modelling mismatch power. The derivation for transition probabilities among KF-bank elements and the production for ultimate output estimate are explained.

- **Chapter Five:** Elaborates on the cooperative ACM by employing IMM-KF feedback link between transmitter and receiver for the purpose of supplying accurate CSI estimates. The transmitter accordingly changes the modulation and coding mappings based on the CSI feedback to arrive at best channel capacity. The water-filling (WF) method as a good candidate for such power adaptation and variations among upper or lower codes is established to enhance the overall transmission throughputs. Imperative CSI parameters such as channel and additive noise power are deliberated.

- **Chapter Six:** Details the conventional AMC methodologies that play a fundamental role in this application and elaborates on the maximum likelihood (ML) algorithms for classification and examines their optimal performance. The development of blind parametric paradigm by exploiting the ML and orchestrated by IMM-KF algorithm for CSI estimate is provided. The EM methodology and comparison to the classical ML estimate (MLE) for CSI interrogation are exercised and the output of which is corroborated against the new paradigm of IMM-KF augmentation. The best fit modulation classes are hence decided which represent the actual constellations received blindly. The performance metric of correct classification rates is also determined in this chapter.

- **Chapter Seven:** Reviews the main groups of classical SS techniques that have a direct impact on the development of this study and develops a novel SS technique that outperforms previous methods and helps achieve better spectral utilisation efficiency. The proposed technique is projected to make use of various analytical and development tools such as spectrograms, filter banks (FBks) and multi-filtering concepts. The MTM will be the central focal for the proposed SS as it has better performance compared to classical methods belonging to the same non-parametric category. This is due to its strict control imposed on the out-of-band (OOB) spectrum leakage and efficiently addresses the bias-variance dilemma. The detection performance metrics of the proposed MTM for standalone SS is

developed to meet the targeted statistical criteria. The adaptive threshold (AT) and optimal frame duration (OFD) by using the newly introduced IMM-KF policy is also investigated in this chapter for further performance gain in terms of spectrum reutilisation and PUs protection requirements.

- **Chapter Eight:** Summarises the overall work analytics and outcomes and induces the main conclusion remarks of this study programme. Further insights into potential future work expansions are also highlighted in this chapter.

Chapter 2

TECHNICAL PRINCIPLES AND REVIEW OF COGNITIVE RADIO

2.1 Introduction

Technology is never dedicated for technology sake; it is mainly adopted to fulfil the good wealth of humanity, anywhere and at any time. An outlet of this aim is also targeted for natural resources and other life kinds on this planet to the best possible extents. CRs come well along this route to enable some of the futuristic major wireless applications. They become a cornerstone in most recent quests aiming at sustainable solutions for the RF spectrum dilemma and thereby meet the hefty demand of broadband wireless pedestals.

The future has been envisioned to be user-centric and have networked societies with unrestricted access to information and sharing of data available anywhere and anytime to anyone and anything [11, 12]. The sheer transition number of connected wireless devices is nowadays expected in the range of 5 to 50 billion. Despite this might appear very exciting, however, it will also pose a formidable challenge for future wireless systems. Present systems mainly operate in spectrum below 6 GHz [10, 11], and going to higher ranges suffer from propagation challenges and therefore other feasible options need to be examined [12]. This can be attained either by featuring existing spectrum access technologies or go beyond and further develop new technologies to meet the above challenges and address specific scenarios and desires. CRNs hence offer the opportunity of having multiple independent networks that operate concurrently within the same area and spectrum and allow for efficient spectrum usage at low traffic costs.

This chapter provides further insight into the radio technology evolution, SDR and CR concepts and standards, and state of the art in such lively field of modern wireless communication systems. The purpose of which is to have a good understanding of the current and new trends and directions in support of the analysis and developments of appropriate dynamic CR techniques delivered in the following chapters.

2.2 Evolution of Cognitive Radio

The history of traditional radios can be traced back in time to the pioneering work of Guglielmo Marconi between 1899 and 1901 when he was the first to experiment on wireless telegraph machines and achieved success in a transatlantic transmission [1, 12]. On the other hand, the concept of CR is still under continuous conceptualisation and development. CR is viewed as a potential driver for making significant changes in accessing and improving the RF spectrum utilisation as one of its primary objectives, and as such CR can be justifiably described as a "disruptive, but unobtrusive technology" [1].

The concept of dynamic or opportunistic spectrum sharing is not a novel concept and is probably as old as radio communication itself [14]. Historically, all traditional radios before 1960s were known to be mostly made of fixed hardware designs and components [15, 16]. The first sign of adding software capabilities in radios was in the 1960s when the FCC allowed the use of shared channels, citizen band, and in land mobile communication systems. With the advent of wireless data communication in the 1970s, the Aloha protocol [14] was proposed to enable sharing of the radio channels for wireless data communication without using a centralised entity. That was the second milestone in embedding more programmable features in radio systems. The FCC then indirectly paved the road to describe the spectrum coexistence procedures for low power wireless devices in the ISM bands by issuing Rule Part 15 in 1985. The first attempt at realizing SDR was in 1987 through a project called Integrated Communications, Navigation, Identification Avionics (ICNIA) by Air Force Rome Labs (AFRL), which was followed by the Defence Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory (AFRL) SPEAKeasy I and II projects in 1990 and 1996, respectively.

In the late 1990s, nearly all radios comprised universal chips of a digital signal processor (DSP) to perform low level modulation/demodulation and signal processing functions, and a general purpose processor (GPP) for high level operator interface, network connection, and system computation functions [16]. Such a major technological step then opened wide doors for the proliferation of aware and adaptive radio systems. The software ingredients were desirably added to radio designs to have

advantageous upgradable capabilities, extended life cycles, reduce cost overheads among others.

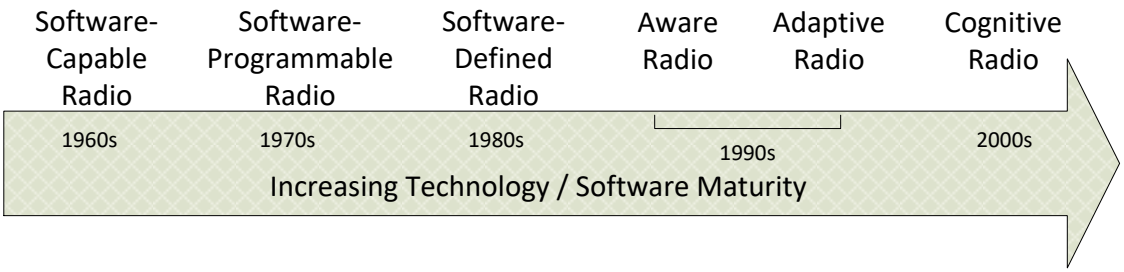


Fig. 2.1 Roadmap of CR evolution.

Later the FCC, and other alliances worldwide such as the International Telecommunications Union (ITU), the Institute of Electrical and Electronics Engineers (IEEE) P1900 study group, and the European Union (EU) End-to-End Re-configurability project (E2R) [2], realised that CR techniques are the future substrate that stimulates the full growth of open spectrum. They have subsequently launched several significant studies on CR to explore its technical and economical capabilities and feasibilities. Such endeavours were capped by the advent of IEEE 802.22 standard, and its equivalence elsewhere, which was released in 2004 to legalise sharing of the TVWS bands [2, 4-8, 17-19]. That initiated the new generation of CR systems in the late 2000s to make many radios benefit of various software classes to enhance adaptation functionalities and significantly support user extensions [16]. Thus, with minimal additional hardware, additional software features were mainly destined to enable users, network operators, spectrum owners and regulators to accomplish much more sophisticated tasks than was possible with the earlier generations of fixed hardware radios.

The above summary on the roadmap of radio technology advances is depicted in Fig. 2.1. Further inception into the evolvement of a wider scale of radio technologies can be consulted in [14-16], at the time of their publicity became available. Before embarking on the state of the art studies on CR systems based on SDR technology, the technical practicalities and working standards of this vivid scientific field are briefly previewed next.

2.3 Cognitive Radio Practices and Standards

As indicated in Chapter One, the striking findings of poor spectrum utilisation coupled with recent advancements in radio technology led frequency regulators to revisit the traditional way of spectrum management. The scarcity of radio resources has become a shocking fact hindering the progression of several wireless applications such as the different generations of mobile communication systems. This negative impact not only influenced broadband access in urban, suburban, rural and remote areas, but also affected other domains such as public safety, health care, business, and leisure among others [17].

Empirical campaigns have vetted that about 15% of households mainly located in rural areas cannot access broadband services in the UK only, while 30% of the rural population has no access to high speed Internet in the EU [4]. The same low coverage figures are also anticipated elsewhere including the USA. Consequently, operators in the UK, EU and USA among others attempted to pursue other wireless possibilities and the TVWS were eventually recognised as a promising opportunity. The proposed network topology could essentially be made of point-to-point or point-to-multipoint terminal connections. That initially pointed to WiFi, line-of-sight (LOS) worldwide interoperability for microwave access (WiMAX), or time-division long term evolution (TD-LTE) as competent air interface technologies besides SDR based CR systems [4, 10, 14]. TVWS are attractive for use with rural broadband for definite reasons. First and for most because it is free and secondly because it is fairly stable; which means that PUs rarely change their bands of use in such areas. Since it is fixed point-to-multipoint, the probability of other SUs causing interference is almost slim to none. This can be feasible subject to have some sort of formidable measures in place to protect against such unlikely events and null this probability.

Under the FCC directions and in recognition of the above mentioned detrimental fact, the IEEE 802.22 family of standards has been developed and industrialised to specifically focus on rural broadband applications [2, 4-8, 17-19]. The emergence of IEEE 802.22 represents a cornerstone in the development of wireless rural area networks (WRANs) [2, 8, 10, 17, 18]. It is mainly aimed at allowing CR systems to share geographically sporadic spectrum bands that were originally allocated for TV broadcast services and whenever they become unused. The IEEE 802.22 is considered

timely and therefore has the potential for global wide applicability to bring a remote broadband access to sparse rural areas with low population densities. That is based on license-exempt and subject to not causing any harmful disruption to the incumbent TV operators and other licensed devices such as low-power wireless microphones [19, 20]. The IEEE 802.22 system is typically limited to one base station (BS) and a set of customer premises equipment (CPEs). The locations of CPEs are fixed at certain locations, while the distances between the BS, CPEs and wireless microphones are either fixed or randomly selected in some scenarios.

In response to the exponential growth of spectrum awareness, the new era of spectrum coexistence has thus been started and various systems' operators and regulators merged to culminate their development efforts and industrial standardisation. For this reason, the innovative DSA concepts, such as spectrum pooling and opportunistic spectrum access (OSA), are gaining momentum in academia and industry [14]. Numerous leading companies and research groups in various universities have started to focus on CRS and applications. The famed annual conference called Dynamic Spectrum Access Networks (DySPAN) has been created by the IEEE in 2005 accordingly [7, 8, 15, 16, 20]. The main intention of this international gathering is to share expert opinions and findings on different spectrum issues and advancements using prominent technologies such as CR systems and CRNs.

In an effort for coordinated CR standardisation works, the IEEE structured the IEEE Standard Coordination Committee 41 (IEEE SCC41) to address all issues related to the deployment of NeXt Generation (xG) radio systems [2, 7] and advanced spectrum management [8, 14, 17, 18]. Then in 2007, the IEEE created a new governing body for all IEEE 1900 standards and named it SCC41 on DSA networks (DSANs). The IEEE SCC41 constituted four working groups (WGs), identified as IEEE 1900.x where (x) represents one of the WGs, and one study group each have been chartered for developing standardisation processes for different aspects of CR-based air interface. Other IEEE projects related to xG radios, like IEEE 802.{11, 15, 16, 19, 20, 21, 22} are knowledge-based sources for IEEE SCC41 [14, 18, 20]. Keeping an actively continuous cooperation between broad standard bodies is crucial during CR developments, the IEEE SCC41 has therefore teamed up with the FCC, Ofcom, SDR Forum, and Object Management Group Forum, to name a few. The internal and external relationships between different WGs of IEEE SCC41 and other standardisation bodies are depicted in

Fig. 2.2. It is worth to mention that despite the IEEE 802.22 is fully dedicated for spectrum access of TVWS, however, the essence of spectrum coexistence and sharing is also readily applicable to other IEEE 802 family of standards, generally in the upper frequency ranges.

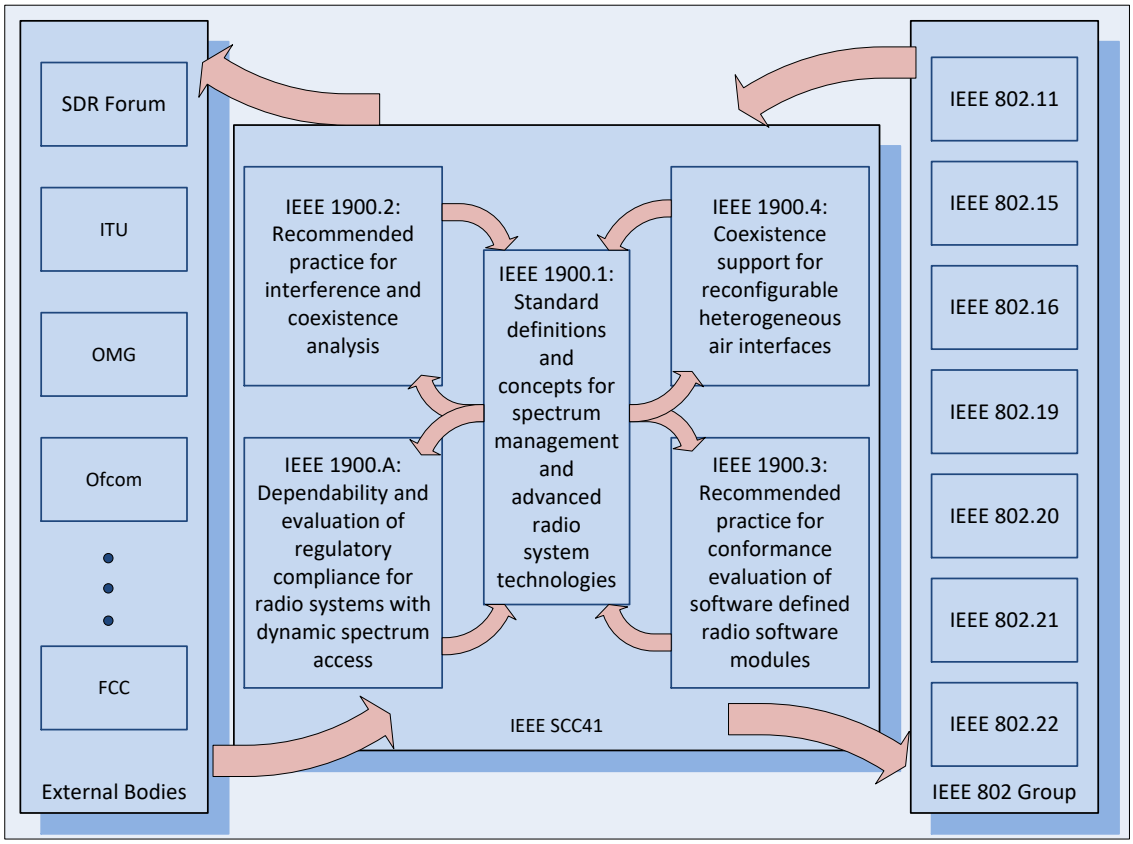


Fig. 2.2 Internal and external IEEE SCC41 interaction relationships.

It can be deduced from Fig. 2.2 that the final endorsements of coexistence among various air interfaces are yet to be consented. This is chiefly because the framework of CR is not restricted to only the radio layer and therefore the final and full definition of CR is still under great debate. Until all parties harmonised consensus on the final definition of CR, an alternative term called cognitive functionality in wireless communication network is temporarily used to represent the art of efficient use of spectrum [14]. This term is aimed at invoking efficient spectrum utilisation by having coordinated processes spread across all layers of the systems' architecture. The execution of this goal is greatly hinged on varieties of technical, policy, and regulatory

issues. Table 2.1 below provides a brief description of the IEEE 802 wireless standards exhibited in Fig. 2.2 above.

Table 2.1 A brief description of the IEEE 802 standards supporting CR, DSA, and coexistence technologies.

Standard	Scope
802.11x	This is for wide-area cellular (Wireless Local Area Network (WLAN)), vehicular and short-range WiFi radios (Wireless Personal Area Network (WPAN)). It is considered as one of the fastest and cheapest broadband wireless access (BWA) systems and an earliest step towards spectrum coexistence. It can be used in the ISM bands 2.4 GHz -to- 5.7 GHz and easily deployed in homes or offices [2, 4, 6, 7, 12, 13, 15, 16, 18].
802.15.x	This is for slow-rate, small-power and short range Ultra-Wide Band (UWB), Bluetooth and ZigBee WLAN and WPAN radios and typically targets applications like machine-to-machine and smart grids [7, 13, 15, 16, 18]. It embraces dynamic channel selection to grapple with coexistence and interference issues with other wireless devices supporting 802.11 protocols [18].
802.16x	This is for WiMAX, the most sophisticated and closest relative to IEEE 802.22. It provides wireless Internet access for stationary and low-mobility users. It was initially intended for 10 GHz -to- 66 GHz NLOS operation bands and then extended to cover 2 GHz -to- 11 GHz LOS bands [8, 12, 16, 18, 19].
802.19	This is for recommendations to specify radio-technology-independent methods for coexistence among dissimilar or independently operated wireless devices and networks to achieve efficient spectrum utilisation [5, 6, 7, 14, 15, 18].
802.20	This is for Mobile BWA, or so called MobileFi, and aims at supporting mobility and multivendor standardisation [20]. Examples include various vehicular mobility classes up to 250 km/h in a Metropolitan Area Network (MAN) environment and targets spectral efficiencies, larger sustained user data rates, and larger numbers of active users [21, 22].
802.21	This is for media-independent handover services. It provides a set of extensible mechanisms targeted to enable the optimisation of handovers between heterogeneous IEEE 802 systems as well as facilitate handovers between IEEE 802 and cellular systems [20, 23].

Fig. 2.3 below portrays the relevant bearing of IEEE 802.22 with respect to other IEEE 802 popular wireless standards developed by the IEEE 802 LAN/MAN SCC in terms of network type, distance range and frequency [8, 15, 17].

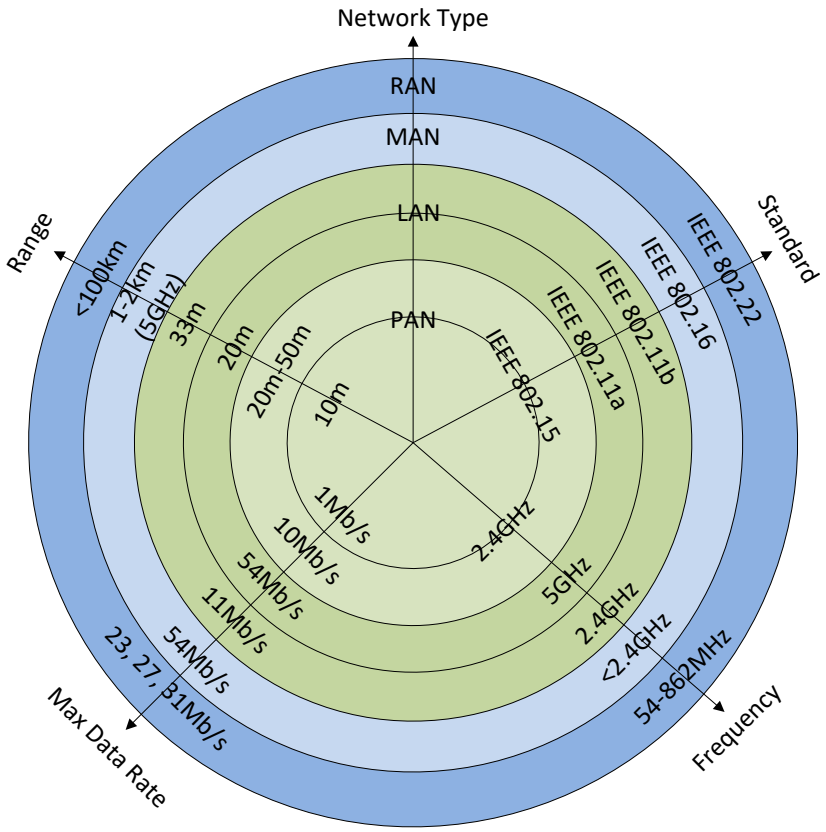


Fig. 2.3 IEEE 802.22 standard relative to other IEEE 802 wireless data transmission standards.

2.4 Software-Defined Radio

As mentioned earlier in the context, earlier attempts on SDR design can be traced back to 1987, when AFRL in the USA funded the military project called ICNIA. That project was a federated design of multiple radios, that is; several single-purpose radios assembled as one piece of bulky equipment [14]. However, the notion of SDR became more explicit when Joseph Mitola in 1991 conducted his PhD work [15] and published earlier findings and proposals in this perspective [24, 25]. The furtherly disseminated results of the same work [26-28] rendered the new SDR concept widely eminent among academic researchers and industrial developers. The most notable statement contributed in Mitola and others' works is that the SDR has been universally seen as a momentous key enabler with vast potential for CR and DSA systems [10, 13-16, 18, 24-28].

The SDR definition refers to a multi-band radio capable of supporting multiple air interfaces and protocols through the use of wideband antennas, RF conversion, and very powerful Analogue to Digital and Digital to Analogue Converters, (ADCs) and

(DACs), respectively. The air interface and all other radio aspects in a typical SDR are implemented on sophisticated GPPs using software routines. Advanced GPPs nowadays are capable of supporting air interfaces that require application specific integrated circuits (ASICs) and make high performance devices and infrastructure practically feasible. The implementation of SDRs is generally doable using a suitable assortment of ASICs, field programmable gate arrays (FPGAs); DSPs in addition to GPPs to produce a reliable telecommunications object [13, 15, 16, 24-28]. A variety of leading SDR projects by academic institutions can be recognised [29-31]. Their generic SDR prototypes may consist of GNU Radio Software and universal software radio peripheral (USRP) for the air interface hardware, as a manifestation of FPGA and other testbed components.

The SDR technology has been evolved to release the radio-based amenities from lingering association with hardwired landscapes and their confinement to frequency channel bandwidths and coding. Hence, the resulting SDR relatively extends the development of programmable hardware platforms used in wireless communication systems and increases their flexibility. The SDR emergence helped improve the QoS by shaping the agility of RF bands, channel admittance forms, data bit rates and errors, and power conciliation. Congruently, the SDR architectures have to attain appropriate trade-offs in terms of the hardware overheads and the rationales of governing policies [24].

Despite SDR heavily relying on digital techniques, however, its divergence from software-controlled digital radio is evident. The crucial variances are mostly accredited to the overall programmability, counting RF, intermediate frequency (IF) making basebands, and modes of channel access and modulation. The supremacy of SDR programmability is largely determined by the use of high performance ADC/DAC technologies embedded amid the IF and channel filtering stages. This would result in the following benefits [24]:

- Enable DSP ahead of signal detection and demodulation;
- Cost reduction of diverse channel access styles by supporting the IF and baseband processing into programmable hardware;
- Focus the component tradeoffs, i.e., provide the computational resources critical to each architecture segment, subject to size, weight, power, and cost constraints.

While for the IF stage, there are wideband digital filters used to select among different service bands. Frequency down-conversion is also performed at this stage in addition to decimation. On the other hand, the signal demodulation, channel estimation, equalisation and decoding are all performed at the baseband stage. The complexity and computational demands of such stages are now easily feasible thanks to powerful microprocessors and digital techniques available in markets. This has led to the design advancement of what is so called the digital front end receivers, which is a major step forward in abandoning their earlier counterparts the conventional analogue front end receivers [32, 33]. This has accordingly furnished a major impetus for greater resilience, consistency and versatility in modern wireless communication systems.

2.5 Cognitive Radio Technology

After the notation of software radio was introduced by Mitola in 1991, he with Gerald Maguire together then firstly adopted the term of CR in 1999 [1, 10, 25]. They both acknowledged the CR as an enhanced extension of SDR and stated the following [25]:

"*Radio etiquette is the set of RF bands, air interfaces, protocols, and spatial and temporal patterns that moderate the use of radio spectrum. Cognitive radio extends the software radio with radio-domain model-based reasoning about such etiquettes*."

The CR based on SDR shortly afterwards has been widely recognised as a thriving intelligent radio that could enhance the flexibility of personal wireless services through a dynamic spectrum sharing and coexistence with other radios of dissimilar platforms [1-20, 24-33]. It instituted for a standard platform that permits a free sharing of frequency pool among many users and able to sustain various radio technologies. This can be approached by unlicensed SUs to perform either overlaying, underlaying or interweaving between unoccupied bands without harming licensed PUs. The fact of having the SS functionality as an indispensable part of CR system to detect spectrum holes and fulfil the DSA objectives has therefore become very tangible.

Wireless communication systems embodying CR aim at two primary objectives [10]:

- High level of reliability anytime and anywhere;
- Efficient frequency spectrum utilisation.

To achieve the above, a CR tasks seamless interaction with its surround to determine appropriate communication parameters and adapt to the dynamic radio environment. The tasks required for adaptive operation in open spectrum are shown in Fig. 2.3, which is referred to as the cognitive cycle (CC) [1, 2, 15, 16, 25, 28]. This cycle is applicable to transmitters and receivers, whether working in tandem or separated.

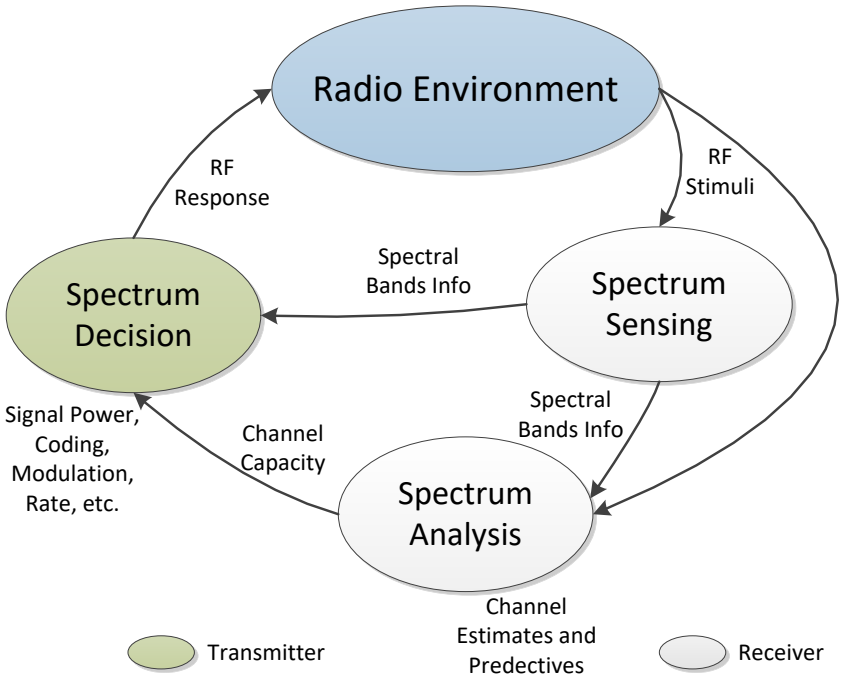


Fig. 2.3 Basic cognitive cycle.

The CC is a real-time enabling process aimed at a CR to interact with its atmosphere, customise rightful transmission parameters, and adhere to the dynamically changing radio settings. It comprises the following three major steps [1, 2, 15, 16]:

4) Spectrum sensing

A CR receiver regularly monitors the available spectrum and acquires the necessary information to detect possible spectrum holes, which covers the following:

- Estimation of interference temperature of the radio environment;

- Detection of unused spectrum holes or bands;
- Estimation of CSI, such as Signal-to-Noise Ratio (SNR);
- Prediction of channel capacity for transmitter attention.

5) Spectrum analysis

The characteristics of spectrum holes that are detected through SS are estimated and analysed, which encompasses the following:

- DSA management;
- Traffic shaping and routing;
- QoS provisioning.

6) Spectrum decision:

A CR decides its own settings after appropriately allocating a spectrum band as per the spectrum characteristics and user requirements, which includes the following:

- Transmit-power control;
- Adaptive modulation and coding, or ACM;
- Transmission rate control.

Once an empty spectrum band is successfully decided, a real-time transmission can be started accordingly over this spectrum band, otherwise it stops until another band becomes available. However, the radio environment is known to be alternating over time and space, a CR is hence to keep continuous monitoring of such changes. Any environmental change during the transmission such as a PU appearance, SU movement, or traffic variations can initiate this adjustment [2]. An overview of meaningful understanding of the interrelationships among CR goals, performance metrics, utility functions, link and network performance, and operating environments can be found in [34]. A firm understanding of these attributes is imperative for a CR designer to proceed effectively.

The implementation of the CR systems with the aforementioned challenging objectives is nowadays feasible thanks to the advancement made in computer software and hardware, DSP, machine learning, and networking. Moreover, the CR is also bestowed with re-configurability in addition to the cognition. That means it can

incorporate new software updates and service applications as they emerge, and diversify development of new interface standards. All the above paved the road towards the SDR platform development and upon which the CR is built, attributed to merging of two elemental technologies: digital radio, and software platforms [1-34].

The SDR technology in a CR system is equally applicable to either or both transmitter and receiver sides, whether in military or commercial environments. As for a transmitter to be cognitive, it has to perform the following intelligent functions [24]:

- Features accessible channels;
- Enquiries a broadcast path;
- Builds an appropriate constellation;
- Automatic beam steering;
- Adjust appropriate power level for transmission.

While for a receiver to be cognitive, it has to perform the following intelligent functions:

- Evaluates energy patterns in the channels of interest;
- Identifies constellations size and mode;
- Adaptively nulls interferences;
- Dynamically estimates the signal and channel properties;
- Adaptively equalises fluctuations;
- Decodes trellis schemes and corrects errors.

From the above, it is obvious that the source cognitive modules need to work in harmony with the sink cognitive modules. Such harmonic cooperation can be usually facilitated by a feedback link channel. This would enhance the receiver capability to deliver performance evidences of the frontward link to the transmitter. However, such cooperative feature between transmitters and receivers is not always visible, especially in unfriendly or military work environments. This would therefore entail advanced automated recognition engines to estimate and determine the radio transmission link parameters as will be seen in the subsequent chapters.

2.6 Cognitive Radio Architecture

There are several varieties of CR architecture and they are all virtually consistent with the initial model proposed earlier [24-27], may differ in minor details, and to have the fundamental CR functions beside other ingredients is their commonality [2, 10, 15, 16, 30, 31, 32, 34, 36]. Mitola described the CR architecture as a framework of integrating evolving products and sequence of designs to synthesise specified functions within particular constraints [35]. He also added that a powerful architecture exploits rapid, cost-effective material and service evolutions. By inspiring all of the earlier views, the proposed functional architecture of a CR system is hence illustrated in Fig. 2.4. This may resemble wearable, mobile, or a radio access point to deploy in a larger heterogeneous network [36].

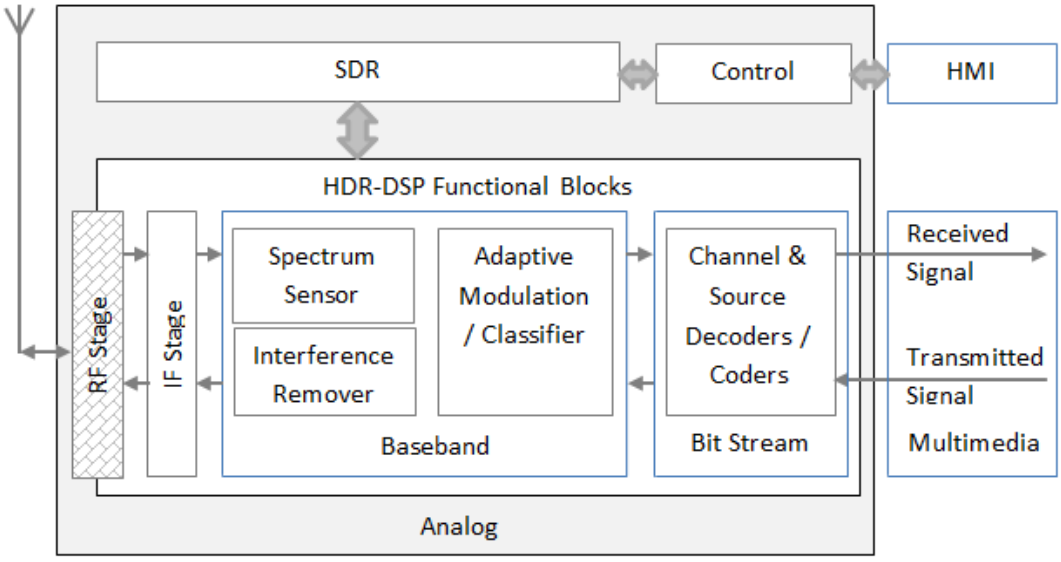


Fig. 2.4 A prototype architecture of CR receiver.

In Fig. 2.4, the RF stage constitutes wideband analogue devices alongside controllable digital components. The digital part of RF stage together with IF stage constitutes the DFE module. All remaining blocks can be operated using DSP functional modules. Various operational and control objectives can be attained through the human machine interface (HMI). The multimedia formats and quality, such as audio-video streams and plugin materials, can thus be altered upon the user's desire and surrounding conditions.

Furthermore, the bit stream and packet pipelining allotted for SDR is achieved as per the CR protocol scheme shown in Fig. 2.5, which is equivalent to the open system interface (OSI) layering architecture [2, 15, 16, 31, 37]. The application layer resides on top of all layers and consists of high-level HMI software and graphics. The SDR layer is equivalent to the PHY layer front-end and performs the actual processing on the transceiver signals. The adaptive engine is performed in the medium access control (MAC), network, transport, and applications layers via especially dedicated protocols. The CR control unit links and provides suitable interface among all modules by using intelligent algorithms to process signals from the PHY layer, and acquires transmission information and requirements from applications to adjust the protocol parameters in different layers.

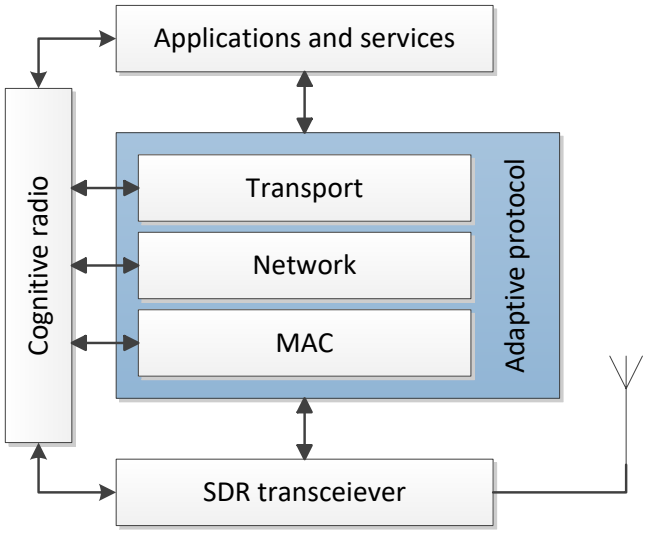


Fig. 2.5 CR protocol stack and OSI layered architecture.

A mix of various programming languages can be employed to reconfigure CR systems and support their desired tasks and functions. For example of higher level languages, the radio knowledge representation language (RKRL) and knowledge query and manipulation language (KQML) can be mentioned [24-28]. The RKRL allows possible sharing of internal SDR system data on the air and be available for specific network purposes using certain interfaces. Hence, a network can pose any unanticipated question using RKRL standard, and any RKRL-capable radio can answer accordingly. The RKRL and KQML are both capable to do internal structure readjustment and customisation. Hence, mobile nodes and networks may share plans about anticipated

needs for spectrum so that it may be efficiently identified and rented. Other structured protocols may be implemented by using the hardware description languages (HDLs), primarily Verilog HDL (VHDL), which readily articulates the internal structure of ASICs and the characters of FPGAs, or an algorithm in the SDR itself [24, 25, 31]. Other lower level interactions may involve C++ and Assembly languages [30].

As mentioned above, the implementation of CRSs is nowadays straightforward due the new advancements made in computer software and hardware, DSP, machine learning, and networking. Also, CR systems are also known to have the features of re-configurability and modularity alongside the cognition. New software updates and service applications can be incorporated in CR systems upon their emergence and the developments of new interface standards can be diversified as well. All that is credited to merging of two elemental technologies: digital radio, and software platforms [24]. The high-end technology progressions of heterogeneous multiprocessing hardware [24, 26] and specifically dedicated system-on-chip [24, 31, 35] made various CR constructions readily achievable.

2.7 State of the Art

This section offers a review of recent literature on the three core CR functions. Some of the earlier notable studies will be explicitly presented or to be referred to indirectly within the context of recent studies as applicable. Specific works will also be addressed in the next chapters to support the analysis and development paradigm of this research program.

2.7.1 Adaptive Coded Modulation

Before delving into the details of AMC conception and provisions, it is important to understand one fact about the TVWS. In the sense of absolutely pure bands, the alleged white spaces are actually not so “white”. They naturally suffer from “pollution” because of low-power signals originated from adjacent DTV apparatuses [38]. Despite these weak signals may be unidentifiable and useless; they are still assumed to be a cause of disturbance for TVWS devices. Nevertheless, this “pollution” does not constitute any significant ruin to the prodigious value of white spaces. That is true since nearly all

modern radio technologies employ ACM techniques to survive very high interference levels and thereby can achieve error-free and spectrally efficient operation.

Generally, the ACM framework is attributed to the nonstop upsurge in massive online multimedia and video streaming services. This has put arduous challenges on existing and future wireless communication designs to be more competent to win the strenuous evolvement race. The new cultural trend of many users worldwide is highly concentrated around the heavy use of modern communication and computing systems and services. Users need unrestricted and reliable access to such applications at anytime and anywhere without any unpleasant discontinuity. This accordingly stresses the ability to have spectrally efficient access and flexible data rates as important design particulars of wireless communication systems. To attend the aforementioned requirements; the need to involve the ACM technology in the design of modern communication systems has become as a prime stipulation. The ACM techniques are normally intended for vigorous spectrum use and vibrant QoS levels to satisfy a wide range of users' requirements and systems' work conditions. Hence, ample of recent research endeavours on coding and modulation were primarily concerned at choosing the best combination of transmission-reception arrangements for efficient resources use in terms of bandwidth, power, and complexity in order to achieve the ultimate reliable performances [39-58].

The inception of ACM notion and rationales is dated long back to the earlier days of telecommunication systems and much before the notion of CR envisaged. As a matter of fact, the literature and history of ACM is very rich and cannot all be grasped here. An insightful broad treatment on the technical foundations of ACM with chief progression milestones can be found in other studies [39, 40], while directions on essential design criteria and trade-offs with a brief history can be found in [41-43]. The information-theoretic approach for ACM outages under MIMO ergodic and non-ergodic capacity definitions is given in [44]. Two major corner stones in the history of ACM are outlined here for the purpose of illustration completion. The first "Big Bang" of ACM plausibility is credited to the pioneering work of Claude Shannon when he exposed in 1948 the relationship between dynamic coding and modulation. He revealed that there exists a coding scheme that achieves an arbitrarily small probability of error for any transmission rate less than or equal to a parameter called "channel capacity", and hence can make a reliable flawless transmission over channel [39-44]. Various hard

attempts since then endeavoured to develop practically feasible ACM implementations with the aim of ideal capacity and performance. The second major breakthrough occurred when Jeremiah Hayes explicitly pointed on the ACM plausibility to mitigate the detrimental channel effects by adaptively adjusting the modulation, channel coding and other system parameters with the aim of a feedback channel as early as 1968 [40].

Successive momentous waves of shift paradigm continued until the multilevel, or M-ary, QAM (MQAM) and the trellis coded modulation (TCM) schemes were invented in 1980s [42]. However, many significant publications on optimum schemes started to appear in the middle of 1990s onward [43]. The MQAM and TCM were initially destined for flat-fading channels with additive white Gaussian noise (AWGN) impairment, but their applications soon extended to cover other types of channels. The main drive behind the evolvement of MQAM and TCM is the increasing demand for higher bandwidth efficiencies, especially for terrestrial radio links, given the constraint of limited bandwidth availability. Generally, the adaptation may include one or a combination of the following variables; rate, power, code, modulation, or bit error rate (BER) among possible others [42, 43]. Moreover, the advent of TCM was a flag on the importance of combining modulation and coding in a single entity for better spectral efficiency, the concept of which was firstly coined by Ungerboeck's pioneering work [42, 45, 46]. As higher power efficiencies were sought, more powerful error-correcting codes were eventually made available.

The fitness of TCM and MQAM adaptation techniques under fixed and fading channels were addressed in many literatures and only few notable are be mentioned here [47-50]. The MQAM, or sometimes named adaptive QAM (AQAM) [40], and with the TCM are both intrinsically tasked to perform link adaptation for high-speed data transmissions. Variable rates and powers have been investigated and the results showed that various considerable gains can be achieved by applying the MQAM technique [47, 48]. In addition to the MQAM gains, the results in [49, 50] also showed that combining the TCM with MQAM produces further benefits compared to the non-adaptive trellis codes. Assume each adaptation variable signifies one degree of freedom (DoF), it was suggested that enough to use one or two DoFs for ACM yields close to the maximum possible spectral efficiency obtained by utilizing all DoFs [51]. Therefore, the parameters to adapt are applications specific and determined by practicality appraisals.

Despite their attractive features under AWGN conditions, TCM schemes do not recognise the rapidly changing wireless variations and their performance could easily become compromised. That is because the TCM performance is highly susceptible to burst errors under severely fading effects, similar to Rayleigh and Rician channels, where extreme mobility is expected [52, 53]. As alleviation to such indulged circumstances, another modified adaptation scheme called bit-interleaved coded modulation (BICM), which is based on powerful convolutional codes, can be employed [52, 53]. BICM enjoys a time diversity of error-correcting codes, which makes it robust against channel prediction errors. This merely leads to negligible performance degradation under the sole provision of marginal channel knowledge based on delayed or outdated estimates, which is the opposite of TCM situation. Adaptive BICM based on turbo convolutional codes is also dubbed turbo-coded modulation (TuCM), and the output of which can be directly mapped onto a high-level modulation of MQAM. TuCM adapts the encoder stream, constellation size and power and hence near capacity throughputs can be readily achieved under severe fading conditions. TCM differs from that as it delivers variable rate of un-coded bits only.

So much as that, it has become apparent that ACM in many respects involves accurate CSI estimates at receivers and feedback essential details to transmitters with minimal latency and error. Albeit some of the above studies did not explicitly specify much on the feedback links, i.e., CSI components were either ignored or assumed perfectly known to the transmitter, the need of well-defined CSI estimators and predictors is a must in practical implementations [43]. The CSI vigilance remarkably enhances the ACM reliability, sensitivity and performance. Therefore, various studies addressed this crucial necessity and committed arbitrary CSI estimators or predictors in their designs, for example the effect of minimum mean square error (MMSE) estimator was quantified in [44, 45], LS, linear MMSE (LMMSE), MLE estimation algorithms were considered in [46], while a quadratic power predictor was employed in [48, 50].

The functional block of channel and source coding/decoding in Fig. 2.4 is assigned to the recovered bit stream to perform conventional channel equalisation, source error correction decoding, frame alignment, bit stuffing, and radio link encryption [24]. Many coding schemes are accessible and of which the hybrid automatic repeat request (HARQ) can be named [54]. The HARQ and AMC are both considered as excellent strategic options commonly used in modern wireless systems to converse over error-

prone and time-varying channels. The list of efficient channel and error-correcting codes can be very long [39, 40]; however, the provision is that their application shall be cooperated among transmitters and receivers to fight uncertainties and errors. This is typical in civil communication platforms; whereas receivers blindly strive to estimate the transmission coding and constellations by means of more advanced algorithms under hostile and unfavourable exposures.

CRs featuring ACM is gradually increasing with time. There is no better rationalisation witnessed today than the authoritative engagement of refined ACM techniques in modern wireless communication systems, and CRs are unconditionally sitting on the peak of such top notch systems. In fact most of the ACM techniques, if not all, can be directly extended for any high-end task or application in CR systems. The need for adaptation with the aim of spectral efficiency enhancement has been well established as pointed out in some of the studies mentioned earlier [1, 2, 9]. The parameters of CR can be amenable for adaptation at the beginning and during each transmission session in accordance to the spectrum and channel characteristics. These parameters are dynamically amended in such away whenever a CR switches to different spectrum bands, the transmitter and receiver parameters are continuously reconfigured and the appropriate communication and modulation protocols are used accordingly. A plethora of ACM research progressed so explosively to invoke the CR features, and only a few are brought forward here to reveal the gravity of such field of study. A proactive SS by closed-loop control schemes, such as power control, ACM, ARQ, and feedback channels have been reported [55]. The PU receiver can thus report the quality of the received signal back to the PU transmitter to allow for transmission parameters adjustment. It was also emphasised that a CR must be able not only to detect whether a PU signal exists but also to identify its kind and an interesting approach is to recognise the modulation and coding scheme of the detected signal [56]. Uncooperative CRN exploits this multilevel recognition sensing as implicit CSI feedback of the PU link in order to constantly monitor the impact of its aggregated interference and maximise the CRN throughput. In a relaying situation among cooperative nodes, the adaptive turbo TCM (TTCM) was conceived for appropriately adjusting both the code rate and the modulation mode according to the near-instantaneous channel conditions between relaying nodes [57]. On a different perspective, the DSA game theory was explored in [58].

2.7.2 Automatic Modulation Classification

The task of AMC has been well addressed in the literature and only a few notable are selected here [59-87]. Very attractive general surveys and tutorials on the classical and new trends in the AMC techniques can be consulted in [59, 60]. A meaningful overview of the recent status of AMC algorithms, including a general description of the ML and feature-based (FB) approaches to further details on the realism of cyclostationary-based features has been instituted in [61]. A brief survey dedicated for AMC using MIMO systems can be accessed in [62], while a helpful tutorial on cyclic feature detection algorithms has been presented in [63]. It is worth to recall that the AMC part in CRs of high performance has been marginally introduced in Chapter One through studies [9, 10]. Having said that, the vivid wave of AMC continues and some recent textbooks, based on classical and evolutionary methods, can also be found in the open literature [64].

Day by day, the solid consistency among ACM, AMC and DSA techniques in general wireless communication systems, and particularly in CRs, is reasserted. That is taking place while most developers and operators keep an eye wide opened towards futuristic mobility scenarios such as emerging 5G applications and spectrum opportunities under inconsistent channels [31, 65-67]. The study in [31] redefined SS as the combination of signal detection and modulation classification and used the general term AMC to denote this combined process. While in another study, the reckoning particulars between ACM and AMC techniques have been reaffirmed [65]. The proposed approach for modelling the dynamic behaviour of realistic time-varying fading channels was premised on the famed discrete-states Markov chain (DSMC) scheme. Then a joint estimation paradigm, which relies on the Bayesian stochastic inference and a maximum a posteriori (MAP) criterion, was developed to obtain time-correlated fading states sequentially, at the same time of classifying unknown modulation patterns. It was demonstrated that, by fully exploiting the underlying dynamics of the estimated fading gain of DSMC channel, the recognition performance of adaptive modulations was shown to possess a considerable improvement. It was alleged that the proposed amalgamation framework of AMC with sequential channel estimation poses great promise to render link adaptations more flexible and effective. In another study, a call for electro hyperspace (beyond frequency and time), not only as

part of CR systems but also as a communication environment awareness component of an adaptive spectrum hyperspace access (ASHA) paradigm beyond DSA, was shouted out very firmly [66]. The ASHA paradigm essentially involves the adaptation of sensing parameters autonomously to ensure robust signal recognition, parameter estimation, and interference avoidance. Such an approach was thereby supposed to take the lead in identifying likely communication opportunities in different dimensions of the spectrum hyperspace. ASHA was also presumed to deliver necessary information about the air interfaces, access techniques, and waveforms that are deployed over the monitored spectrum to achieve better resources and interference management. In the same manner, the term Radiobot was newly suggested to refer to the elegant mixture of SS, ACM and AMC features in CRs [67].

The modulation classifier is generally an intermediate step between signal detection and demodulation [59, 68]. The key purpose of a classifier is presumed to correctly choose the modulation format of the incoming signal from a pool of candidate modulations, or to decide that the modulation format cannot be recognised. A desirable classifier should provide a high probability of correct classification in a short observation time interval, especially when the variation in the SNR value is considerably large. It should also have the capability to recognise various modulations under diverse propagation characteristics, model-mismatch robustness, real-time performance, and low computational demand with limited or no prior knowledge about the parameters of the signals. It is a tedious task in non-cooperative environment such as; multipath propagation, frequency-selectivity and time-varying channel, in addition to no prior knowledge of the incoming signal.

Two steps are involved in the design of modulation classifier; signal pre-processing and adequate classification algorithm. Pre-processing tasks may include, but not limited to, noise reduction, estimation of carrier frequency and signal power, and channel equalisation. While on the other hand, there are two main modules of classification algorithms; LB and FB, the second term alternatively named as pattern recognition in some literature [9, 59-72]. The output of both the LB and FB algorithms is the result of minimisation of the probability of false classification. The complexity of such algorithms usually gives rise to suboptimal classifiers. While on the other hand, several features are usually employed in FB classifiers and their observations determine the desired decision. Once the identification of the received signal has been successfully

accomplished, the signal demodulation and information extraction can then be readily performed. Depending on the chosen model for the unknown parameters, three LB-AMC techniques can be proposed; average likelihood ratio test (ALRT), general likelihood ratio test (GLRT), and hybrid likelihood ratio test (HLRT), in addition to quasi ALRT and quasi HLRT. Subject to trade-off between complexity and performance, the quasi-HLRT (or QHLRT) classifiers have generally shown acceptable results without extra burdens.

The blind recognition of modulated signals is often considered as labour intensive in both open and restricted communication systems. The criticality of blindness classification has become stimulating and attracted much attention recently [66-72]. That is due to many unknown parameters in addition to complex channels, such as carrier frequency, symbol timing, etc., that have to be extracted from the received signal. In military and tactical communication systems, such as electronic surveillance, interference identification, and jamming signal design, etc., advanced techniques of real-time signal interception and processing are vital for decisions involving electronic warfare operations and other special actions. This has eventually materialised the determination for smart receivers that are equipped with blind signal processing techniques. It has been shown that the statistical properties of correlation functions of the received signals exhibit peaks at a particular set of time lags for certain modulation formats [69]. Such peaks can be exploited as a discriminating feature in the blind AMC algorithm over single and multiple antennas scenarios. The FB-AMC was applied and the effectiveness of which examined using four different features; the classification tree, K-nearest neighbours (KNNs), artificial neural networks (ANNs), and support vector machines (SVMs) in [70]. The blind signal modulation recognition using clustering analysis and featuring constellation signature was discussed in [71], while the non-parametric LB-AMC using minimum distance centroid estimator was addressed in [72].

In the literature on the LB-AMC techniques based on ML methodology, the following can be deliberated [59-62, 64-66, 72-78]. In this method, the AMC problem is modelled as a multi-hypotheses test (MHT) and the likelihood functions (LFs) of the received signal given the modulation formats are compared against each other. The modulation candidate with the maximum probability density function (PDF) is to be selected accordingly. Such a statistical decision approach is intrinsic to the premise of MAP in detection theory, and from which the ML is induced in case all modulation

candidates have the same probability of occurrence. The PDFs to have all governing statistical variables provided or known in advance is usually not possible and hence solving the multidimensional integrals tangled in the derivation of averaged log likelihood ratio (LLR) tests makes it unfeasible. The unknown variables are usually treated as unknown deterministic values in the classical theory of statistical decision and hence the optimal solution uniformly most powerful (UMP) can be found independent of the unknown variables. However, because of the complicated nature of unknown variables, the UMP test rarely exists in practice. As that is valid for the ALRT, in order to combat this realistic challenge, the MLE of unknown parameters has been developed to substitute in the LFs and build the GLRT algorithm instead [65, 74, 75, 77]. Other estimates are also permissible [65, 72, 75] and hence produce the HLRT or QHLRT which are most common in practice as the averaging is over the unknown symbols only.

On the other hand, FB-AMC techniques have been well addressed in the literature [9, 59-64, 66-72]. Among many features, the wavelet analysis, signal cyclostationary features such as amplitude, phase, frequency, correlation, constellation signature, centroid, in additional to high order statistics (HOS) like method of moments (MoM), cumulants and kurtosis, to name just a few that can be easily discernible in the literature. As for decision making exploiting the above and other features, dozens of algorithms can be employed and of which the ANN, SVM, classification tree, principal component analysis (PCA), independent component analysis (ICA), hidden Markov model (HMM), fuzzy logic and the list goes on. It is advisable to refer to the survey studies for a complete list of FB-AMC techniques, most comprehensive in [59, 60] and partially in [61-63]. Various other classification forms are also reachable in the literature (some covered in [64]); such as Dirichlet process mixture model [67], hierarchical polynomial [79], nonparametric statistical goodness-of-fit for PDF tests like Kolmogorov-Smirnov (K-S) [80], Gabor filter in conjunction with HMM and genetic algorithm (GA) [81], particle filter [82] and Gibbs sampling applied to a latent Dirichlet Bayesian network [83], while evolutionary intelligence algorithms are plenty and may include gravitational search algorithm, teaching-learning-based optimisation, particle swarm optimisation, ant colony optimisation, artificial bee colony, cohort intelligence, adaptive cohort intelligence and the list goes indefinite [84, 85].

Bearing all that in mind, it is worthy to emphasise that any of the above given AMC paradigms are applicable, whether locally or as a collaborative member, within a group of distributed nodes or sensors. It has been shown that such cognitive centralised processing scenarios (multi-sensor AMC) have many advantages in comparison to a single-sensor AMC, such as signal enhancement via diversity, heightened performance and reliability and reduction in the computational complexity and communication cost [86-88]. The main challenge in such scenarios is that all collaborators need to orchestrate their operations, and that could be attained by having a single core or multiple fusion centres.

It is important to understand that a variety of solitary or joint algorithms can be employed using any of the above AMC methods; however, there are no uniquely accurate templates for a perfect signal observation. Due to computational burdens, some of the classification techniques cannot be implemented in real time situations despite their outstanding accuracy. A practical classifier should always retain a good balance between the desired performance accuracy, processing speed and implementation complexity.

2.7.3 Spectrum Sensing

SS represents one of the main engines in CR systems and has been researched extensively in the last few decades, where only a few stylish studies are enumerated here [89-106]. It was stressed in Chapter One that SS involves continuous scanning the RF spectrum and processing the findings to determine which channels (holes or bands instead of channels can be used interchangeably) are free or engaged by PUs. The SS feature is encompassed as mandatory within IEEE 802.22 standard [1, 2, 6, 8, 9]. Two options were discussed in concern of performing the task of finding spectrum holes; the first option is by having a receiver that is equipped to scan a wide frequency range; and the second one is by employing a FBk covering the whole range [14]. Without doubt, the second option is practically not feasible regarding cost and complexity and therefore scanning is the first choice. The pitfall of complete scanning is that observing large spectrum chunks takes a considerable time and the retrieved information is of a statistical nature. Nevertheless, such factual detail does not constitute a major challenge if the scanning procedure is done in a systematic way, as follows. As long as the CR is

not active, it scans the frequency range of interest, evaluates the observations, and saves the results about the spectrum occupation in a local statistical database. Whenever the CR wants to set up a connection, it looks into the database to identify a spectrum hole that is very likely idle. Then it conducts a quick passive measurement to verify if this hole is effectively free. If the new measurement yields an idle hole, this is taken for transmission; if not, the resource with the next priority is taken from the database and verified. It is worth noting that such systematic scanning process was termed as wake, observing, and sleep (prayer) epochs or cycles [1, 35].

The aforesaid studies offer a concise outline on the SS techniques and relevant citations, and one more study can be added to this list [89]. A significant number of studies with thorough treatment and survey elaborations on SS are available in the literature, and only the most eminent are registered here [2, 55, 90-102]. Some insightful textbooks were also publicised recently for further knowledge expansion in this regard [103-106]. Based on all these studies, the global classification of SS techniques can be put under two main categories; non-cooperative or local and cooperative centralised or distributed. The second level of SS categorisation can be broadly put under three main groups: matched filter (MF), energy detector (ED) and cyclostationary detector (CSD), as depicted in Fig. 2.6 below.

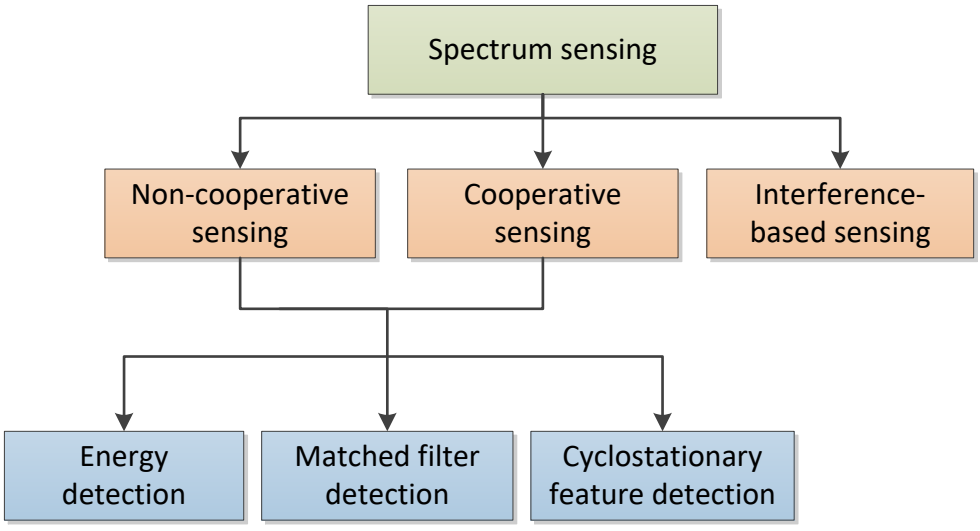


Fig. 2.6 Different types of SS in the CR PHY layer.

The three long-standing SS techniques are briefly explained as follows [2, 55, 90-106].

1. Energy detection:

   The method of ED, also known as radiometry, is the most common SS technique because of its low computational and implementation complexities. In addition, it is more generic as receivers do not need any knowledge on the PUs signal. The signal is detected by comparing the output of an ED with a threshold relevant to the noise floor. Some of the challenges with ED based sensing include selection of the threshold for detecting PUs, inability to distinguish interference from PUs and noise, and poor performance under low SNR values. Moreover, ED does not work efficiently for detecting spread spectrum signals.

2. Matched filter detection:

   The method of MF is perceived to be optimum for detecting PUs when transmitted signals are known. The main advantage of MF is the short time to achieve a certain probability of false alarm or probability of miss-detection as compared to other methods. In fact, the required number of samples grows as SNR values get low for a targeted probability of false alarm under MF approach. Also, MF entails CRs to demodulate received signals. Therefore, it requires a perfect knowledge of PUs signalling features such as bandwidth, operating frequency, modulation type and order, pulse shaping, and frame format. Moreover, since CR needs receivers for all signal types, the implementation complexity of MF sensing unit is impractically large. Other MF difficulties are large power consumption as various algorithms need to be executed for detection and it needs perfect coherence for operation.

3. Cyclostationarity feature detection:

   The method of CSD is employed for detecting PUs' transmissions by exploiting the cyclostationarity features of received signals. These features are caused by the periodicity in signal or in its statistics like mean and autocorrelation, or they can be intentionally induced to assist SS. Instead of power spectral density (PSD), cyclic correlation function is used for detecting signals present in a given spectrum. The CSD algorithms can differentiate noise from PUs' signals. This is

due to the fact that noise is wide-sense stationary (WSS) with no correlation, while modulated signals are cyclostationary in their spectral correlation given the redundancy of signal periodicities. Furthermore, CSDs can be used to distinguish among different types of transmissions and PUs.

New spectrum detection schemes have also been investigated recently; such as wavelets, autocorrelation, covariance, eigenvalue, compressive sensing, blind source separation (BSS), etc. In practice, however, a combination of one or two detection schemes may be augmented to work together and tackle different situations. For example, a swift but coarse scan may be performed using ED to identify a few possibly free bands. While the fine revelation of most suitable candidate bands may then be conducted using a more accurate detection method.

As for the interference-based technique, the sensing algorithms measure interference noise level at PU receiver [1, 2, 91, 93, 94, 97, 98, 103]. The generated information is thus used by SU to compute the expected interference level and control the spectrum access without violating the interference temperature limit. Alternatively, an unlicensed transmitter may observe the feedback signal from a licensed receiver to gain knowledge on the interference level. The interference temperature was attributed to the FCC spectrum policy task force. That team has recommended a paradigm shift in interference assessment to rely more on adaptation and tuning between both the transmitter and receiver, rather than largely fixed operations on the transmitter side alone. The recommendation has created a new measurement unit called the interference temperature, which is intended to quantify and manage the sources of interference in a radio environment. Moreover, the interference temperature limit specifies a “worst case” characterisation of the RF environment in a particular frequency band and at a particular geographic location, where the receiver performance is expected adequately.

Without loss of generality, the realisation of SS techniques broadly involves two modes of operation; namely, reactive or proactive modes, depending on the way CR systems search for empty bands [55, 90, 91, 94, 95]. A CR in the reactive mode operates on an on-demand basis and starts to sense the spectrum only when it has some data to transmit. While a CR in the proactive mode, aims at minimising the delay incurred in finding an idle band by maintaining a list of one or more licensed bands currently available for opportunistic access through periodic sensing of the spectrum.

Of course, the enhanced data transmission responsiveness in proactive SS comes at the expense of elevated computation overhead and longer latency. A tradeoff between rate and reliability must be eventually optimised.

Moreover, the SS framework can be applied using two different operation scenarios; cooperative (ad-hoc networks) and or non-cooperative (stand-alone) [1, 2, 6, 9, 55, 90-106]. While the non-cooperative situation is self-descriptive, however, the cooperative one needs some discussion. The cooperative scenario is adopted when the SS performance is degraded by presence of large noise, shadowing, and multi-path fading effects. When PUs signals are completely immersed in noise background, there would be certain levels of SNR thresholds below which reliable spectrum detection is impossible even with a very long sensing time. If SUs cannot detect the primary transmitter, while the primary receiver is within the SU's transmission range, a hidden PU problem will occur, and the PU transmission will be interfered. In situations of spatial and multiuser (MU) diversity, cooperative SS is proposed to improve the reliability and efficiency of SS, increasing the probability correct detection, and reducing the probability of false alarm. Decision fusion rule is common in centralised cooperative SS. As such, a central controller, e.g., a secondary base station, collects local observations from multiple SUs, decides on the available spectrum channels, and informs the SUs on a channel's access availability. In a reduced cost scenario, SUs in distributed cooperative SS exchange their local detection results among themselves without requiring a backbone infrastructure. Relays can also be used in cooperative SS where the cognitive users operating in the same band help each other relay information using amplify-and-forward protocol. Agility could also be improved in case of network asymmetry.

Despite all the above, there are concerns engulfing the cooperative SS. For instance, ad hoc networks may involve low-class commercial SUs equipped with low-cost and limited power devices and hence cannot afford vastly sophisticated detection demands. In a wideband cooperative SS, multiple SUs have to scan a wide range of spectrum channels leading to large amounts of sensory data exchange and energy consumption in addition to inefficient data throughputs. Last but not least, sensing under vague dynamic situations, like users' mobility and severe channel fading, is considered even more puzzling.

It has become apparent that arriving at an optimal SS policy that meets the coverage and transmission performance settings of various CR systems and networks is crucial. In a rather different approach, an alternative SS scheme based on the non-parametric MTM was favourably voiced for an attractive opinion [1, 90-95, 103, 106]. The particular ingredient of discrete prolate spheroidal sequence (DPSS) wave function as prototype filter coefficients makes the MTM very appealing. These functions, which are also known as Slepian sequence, were characterised by two important properties; 1) they have maximal energy focused in the main lobe, and 2) they are orthonormal. Thereby, they employ orthogonal data sets to achieve balanced tradeoffs in the bias-variance dilemma to enhance the spectral estimate resolution. The SS based on MTM has an outstanding performance unlike any other in the same family, which was attributed to the robust leakage control in the OOB spectrum manifested by the tapering process.

But how can the spectrum be actually accessed and SUs hence survive the coexistence with PUs without causing any damage? The answer to this question is illustrated in Fig. 2.7 below and the details of which are given underneath.

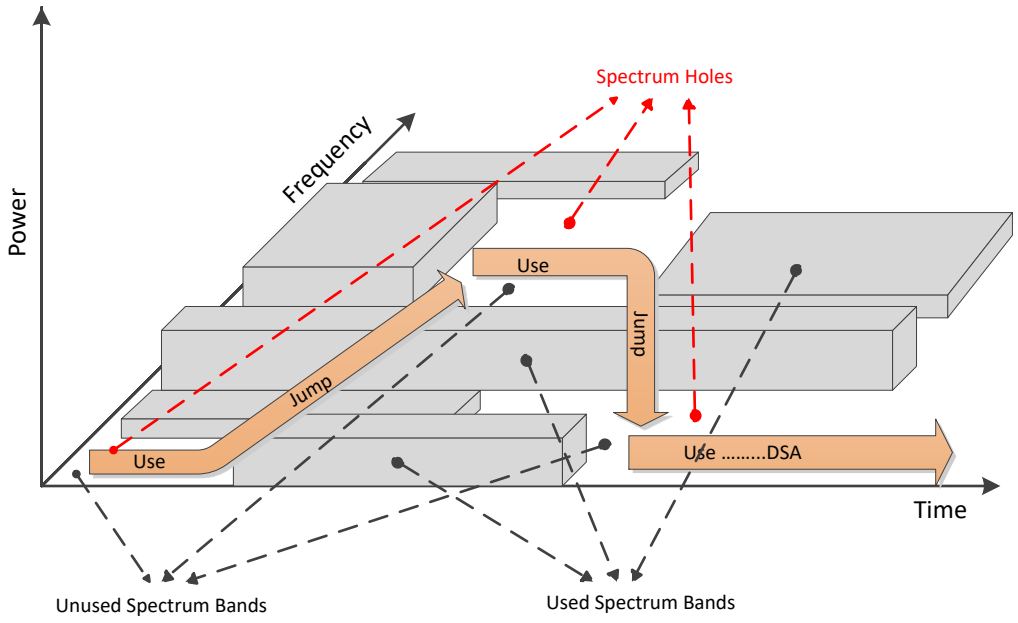


Fig. 2.7 Illustration of spectrum holes and interweaving access.

The DSA technology adopted and implemented in SUs governs the exploitation of opportunistic spectrum bands and coexistence with PUs. Based on that, the existing spectrum exploitation can be broadly categorised into three main groups; 1) interweave, 2) underlay, and 3) overlay techniques [55, 91, 92, 94, 95, 97, 98, 102, 104]. These terminologies are occasionally arguable by different researchers, however, they represent the most common terms used in the open literature. There is also one more kind called the "hybrid" technique, which stands for the mixture between interweaving and overlay spectrum detection and access. The interweave paradigm is based on interference avoidance and opportunistic techniques. SUs in this case communicate deviously using vacant spectral holes in space, frequency, and time and ideally there occurs no interference to PUs. The overlay paradigm is characterised by the mitigation of interference using advanced coding and transmission strategies at the SUs, such as spread spectrum and dirty paper. SUs hence can transmit simultaneously with PUs and the interference caused can be offset by using a part of SUs power to relay PUs. While in the underlay paradigm, the communication is characterised by firm restrictions on the interference that may be caused by SUs to PUs. In this case, the coexistence of PUs and SUs is permissible subject to having interference temperature below certain limits and can be controlled and managed properly.

2.8 Summary

This chapter has elaborated on the SDR and CR concepts, platforms, supporting standards and major corner stones in the emergence and progression of their associated technologies. The core engines of paradigm CRs, which constitute of ACM, AMC, and SS have been expressively surveyed and enlightened. Also the opportunistic access spectrum and coexistence paradigms have been demonstrated. It became apparent with time that CR systems are fuelling the revolutionary eminence of versatile new platforms and technology to push the boundaries of high mobility service even higher than anticipated before. That was obvious by the efficient exploitation of the scarce resources such the RF spectrum and retain proper control and management as such. There left no doubt that CRs will play a key supportive role in the forthcoming advanced ubiquity systems everywhere and at any time.

Chapter 3

WIRELESS TRANSMISSION AND FADING CHANNELS

3.1 Introduction

The modelling of reliable and feasible wireless communication systems is considered as a critical assessment step before real design and implementation can take place. This is important to save unnecessarily expenditure in making tedious adjustments afterwards. It is also equally important to have a full understanding and perception on the nature of wireless channels and their behaviour in the presence of different signalling techniques conveyed over noisy and fading impairments. This chapter is hence intended to provide insightful information on the prevalent wireless transmission techniques and channels. An innovative information-theoretic approach using the Tauchen model borrowed from the economic studies is proposed for characterizing the FSMC model of Rayleigh fading channels, which manifests as the main contribution of the work in this chapter. Simulation exercises are conducted with the aim to validate and verify the given analysis of wireless transmission and the newly developed FSMC channel modelling based on the Tauchen approach.

3.2 Wireless Communication System

This work is concerned with the transmission and reception of linearly modulated digital signals through wireless channels. Different versions of digital wireless communication systems can be found in the literature, the ingredient components of which are application specific [39, 107]. Fig. 3.1 depicts the building diagram of a typical wireless communication system adopted here. The RF, ADC and DAC conversion stages are omitted for the purpose of presentation ease, keeping in mind that the analysis and development of this work are mainly aligned with baseband signals only.

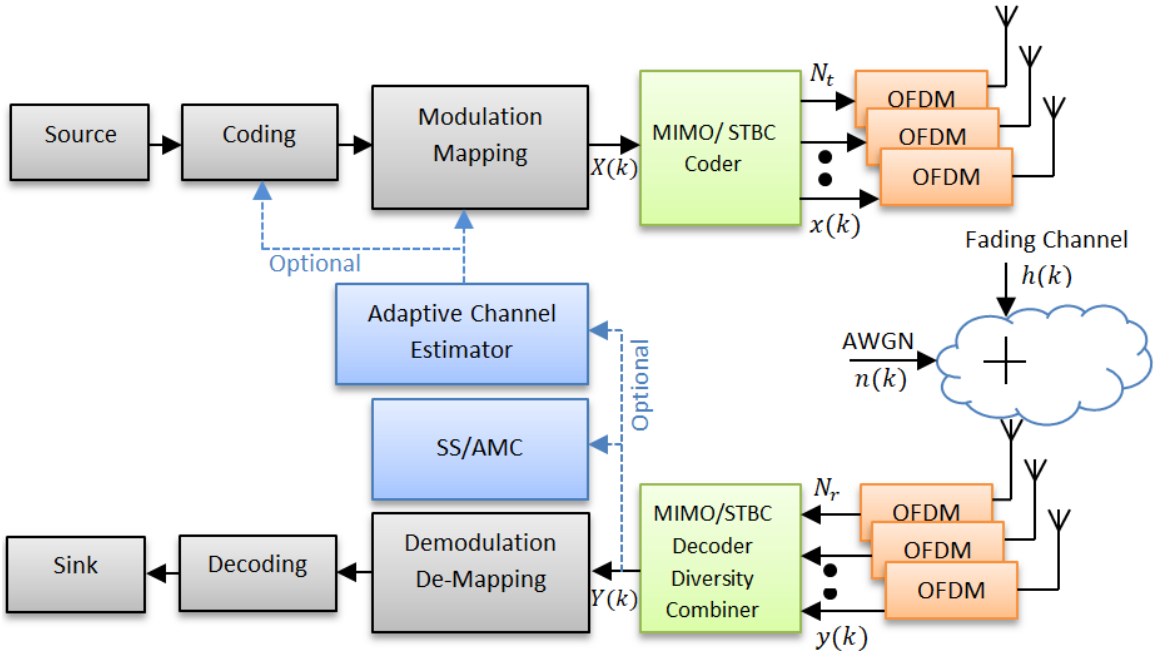


Fig. 3.1 Block diagram of a typical wireless communication system with CR features.

The diagram of Fig. 3.1 is slightly more elaborate than those in the literature as it constitutes the essential components for wireless transmission and reception (given in black and grey colours) besides extra optional components (shown in different colours). The CR functional components (given in blue colour) will be detailed in the subsequent chapters, while the essential and other optional components are explained in what follows. Before delving further, it is necessary to realise a few important facts on digital communication systems. These systems are more robust compared to their analogue counterparts, and all systems allocated for wireless applications are hence digitised.

This is also the main idea behind exploiting the SDR platforms in CR systems to retain handling of the digital signals in a more flexible and straightforward way. Additional advantages related to digital communication over analogue communication importantly include faster data rates, higher spectral and power efficiencies, better immunity to channel defects, sophisticated source and channel coding, better privacy and security, versatile multiple access and diversity strategies, and low construction costs [39, 107]. Often some of these entities are conflicting with each other, and the final choice is based on a trade-off between these objective requirements.

Fundamentally, the task of digital communication is to transfer a digital message from the source to the destination. A digital message is nothing but an ordered sequence of symbols produced by a discrete information source. If there exist only two possible discrete states, then this quantity is referred to as an information bit stream. The bits are binary digits, taking the values of either 1 or 0. In broad transmission scenarios, channel bandwidth sets upper limits to the signalling rate, and noise causes errors to appear in the recovered message. Thus, signalling rate and error probability play central roles in the design and specification of most commercial and strategic digital communication systems.

3.2.1 Signal Model

Assume a signal of interest is conveyed over a wireless channel using arbitrary multiple antennas. The signal is of discrete style and the bit sequence of which spans K samples, which is also a manifestation of the time index. Whether for the ACM, AMC or SS pre-processing purposes in CR settings, the baseband signal over single-input single-output (SISO) antennas can be modelled as follows [2, 39, 42, 55, 74, 93, 103, 104, 105, 107]

$$y(k) = h(k)x(k) + n(k) \tag{3.1}$$

and simplifying by making the time index k implicit yields

$$\boldsymbol{y} = h\boldsymbol{x} + \boldsymbol{n}, \quad \forall k \in [1, K] \tag{3.2}$$

where $\boldsymbol{y} = [y(1)\ y(2)\ \ldots\ y(K)]^{\mathrm{T}}$ denotes the received complex-valued signal observed over K samples, $\boldsymbol{x} = [x(1)\ x(2)\ \ldots\ x(K)]^{\mathrm{T}}$ denotes the transmitted signal,

which could be of any appropriate digital format, the channel effect is $h = ae^{j\theta}$ where a and θ represent unknown amplitude phase variations, respectively, and $\boldsymbol{n} = [n(1)\ n(2)\ \ldots\ n(K)]^{\mathrm{T}}$ represents AWGN of zero-mean and unknown variance attributes. It is important to realise that the time index can be discarded in $h(k)$ and the parameter h can be kept alone in (3.1) and (3.2) if the channel has no rapid fluctuations over the entire signal window of concern. In such case the channel fading type can be called as slow, flat or block.

Suppose a signal is allocated for transmission using MIMO antennas, N_t on transmitter and N_r on receiver. The signal model is thus revised as below [39, 44, 46, 55, 62, 105, 107]

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{N}, \quad \forall k \in [1, K] \tag{3.3}$$

where $\boldsymbol{X}$ is the $N_t \times 1$ data code vector, $\boldsymbol{H}$ is the quasi-static $N_r \times N_t$ channel response matrix and $\boldsymbol{N}$ is the $N_r \times 1$ AWGN noise vector. This noise statistics are characterised as independent and identically distributed (*iid*) zero-mean circularly symmetrical complex Gaussian (ZMCSCG) and denoted by $\mathcal{CN}\sim(0, \sigma_n^2 \boldsymbol{I}_{N_r})$, and $\boldsymbol{I}$ is the identity matrix of order N_r. The $N_r \times N_t$ quasi-static channel coefficient matrix is given by

$$\boldsymbol{H} = \begin{bmatrix} h_{1,1} & h_{1,2} & \ldots & h_{1,N_t} \\ h_{2,1} & h_{2,2} & \ldots & h_{2,N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r,1} & h_{i,j} & \ldots & h_{N_r,N_t} \end{bmatrix} \tag{3.4}$$

3.2.2 Digital Modulation

Unlike their analogue counterparts, digital signals can be generated, transmitted and retrieved very straightforwardly. They are highly appreciated for their processing flexibility and immunity against channels' imperfections. Generally, a digital modulator on the transmitter side involves the function of encoding (so called mapping) a bit stream of finite length onto one of several possible discrete signal patterns, or constellations, ready for transmission. On the other receiver side, a demodulator should be capable of intuitively minimising the probability of detection error by decoding (also called de-mapping) the incoming sequence into its best match in a set of possible patterns pertinent to the transmitted signal. That is true in as much as both transmitter

and receiver recognise each other and work coherently. If such a scenario is not applicable, the receiver should then be able to identify the signal constellation before performing the demodulation process.

There are two main types of digital modulations; linear and nonlinear [39, 107]. The linear type involves embedding the information bits into the amplitude or phase of the transmitted signal while the frequency is kept constant. The nonlinear type, on the other hand, comprises embedding the information bits into the frequency of the transmitted signal while keeping the amplitude and phase constant. Linear modulation generally has better spectral properties compared to nonlinear modulation, as the latter commonly leads to unfavourable spectrum broadening. Therefore, a trade-off is to be exercised to choose between the two modulation methods; however, the linear modulation has a preference over nonlinear modulation in terms of having better power efficiency and resistance against channel variations. The constellation size is to be chosen once the type of modulation is decided. Virtually, the larger constellation size, the higher data rates are achieved, albeit at the cost of more susceptible exposure to noise, fading and devices flaws.

Table 3.1 Constellation points of common digital modulation schemes.

Modulation	**Constellation Point**
MPSK	$C_i^M = e^{-j\theta_i}$, where $\theta_i \in \left\{\frac{2\pi}{M}(i-1), i = 0,1,2,\dots,M-1\right\}$
MQAM	$C_i^M = A_i + jB_i$, where $A_i, B_i \in \left\{2i - 1 - \sqrt{M}, i = 1,2,\dots,\sqrt{M}\right\}$

This study primarily focuses on linear modulation, albeit the given analysis can be readily extended to the nonlinear type as well. Along this way, the renowned MPSK and MQAM, or sometimes called as amplitude-phase keying (APK), techniques are widely adopted in most wireless systems, especially those intended for CR applications, as was indicated earlier in Chapters One and Two. These two M-ary digital modulations of size M alphabets are described in Table 3.1 [39, 107], where the transmitted signal x can take on any i^{th} point in the *iid* constellation space C_i^M. The complex envelope of received signal can be expressed using the in-phase and quadrature constellation and noise terms, as below

$$y_i = \sqrt{y_{I,i}^2 + y_{Q,i}^2} \quad , \quad \varphi = tan^{-1}(y_{Q,i}/y_{I,i}) \tag{3.5}$$

where $y_{I,i} = A_i + n_I$ and $y_{Q,i} = B_i + n_Q$, and each noise component is now having the power of $\sigma_n^2/2$ assuming that they possess *iid* statistics under certain conditions. Given proper time coordination, these digital signals can be received either by featuring an MF or a quadrature correlator aimed for maximising the energy upon detection.

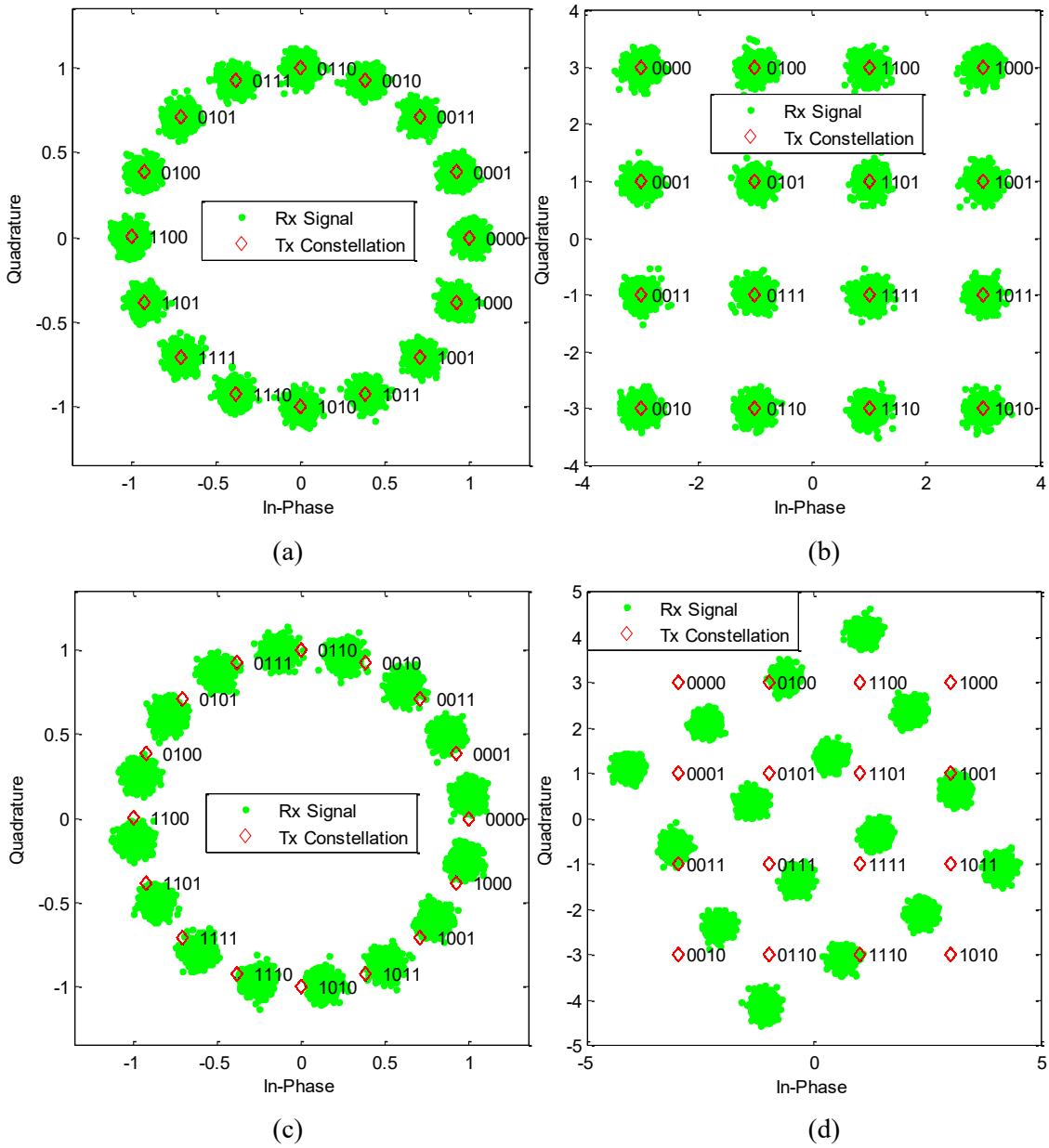


Fig. 3.2 Scatter plots of 16-PSK and 16-QAM (a) and (b) without fading (c) and (d) with fading.

The results of transmitting (Tx) arbitrary selected 16-PSK and 16-QAM signals through an AWGN channel with SNR of 20 dB are depicted in Fig. 3.1. Parts (a) and (b) of this figure show the received (Rx) noisy constellation samples with respect to the original signal using Gray coding and without noise. Due to its attractive merits of not producing propagation errors, Gray coding is very common in mapping of digital constellations. That is, a solitary bit error results due to a single constellation point when mistaken for one of its nearest neighbours [39, 107]. The original signal lies in the centre of received noisy samples is obvious in the above figure, yet the main features of each constellation pattern are still well discerned. Also the distances, or gaps, between adjacent constellation points are distinctly recognised and hence the demodulation process can be applied without hindrances.

Conversely, the above situation looks entirely different when the same signal is passed through a fading channel. Parts (c) and (d) of Fig. 3.2 illustrate the effect of unity amplitude and 30^{o} fading channel on the received signal for the same aforesaid signalling schemes. Dramatic changes occur to the locations of constellation points and their inner spaces, rendering the demodulation process near impossible without having supportive aids. The received constellation points are now entirely deformed and not coinciding with their original locations and the decision of accurate modulation type is thereby very compelling. This is the main drive behind why the AMC and adaptive CSI estimation techniques are unavoidable to combat fading channels, achieve efficient CSI estimation and ultimately lead to a successful AMC.

3.2.3 Transmission Design Characteristics

It has been shown in the context of deciding which digital signalling to be used for transmission is a tradeoff determined by several design factors, such as; transmission rate, error probability, spectral (or bandwidth) efficiency, power efficiency, spectral spill over, and hardware requirements [39, 42, 107]. A prerequisite comparison between contending signalling schemes is hence meaningful to evaluate the designated overall system performance under known or anticipated constraints. Having said that, certain terms are to be defined which establish the basis for comparison between digital modulation systems.

The SNR is on top in the list of several other strategic design parameters. It represents the ratio of received signal power σ_x^2 to the noise power σ_n^2 , where both occupy the same transmission bandwidth B. This is given as per the following general expression, taking into account the engagement of MIMO configuration and fading channel effect

$$\text{SNR} = \gamma_s = \frac{\mathbb{E}[\text{Tr}(\boldsymbol{HXX}^*\boldsymbol{H}^*)]}{\mathbb{E}[\text{Tr}(\boldsymbol{NN}^*)]} \tag{3.6}$$

where γ_s is the signal power to noise power ratio, $\mathbb{E}[\text{Tr}(\boldsymbol{XX}^*)] = \mathbb{E}[\|\boldsymbol{X}\|^2]$ is the signal power, $\mathbb{E}[\text{Tr}(\boldsymbol{HH}^*)] = \mathbb{E}[\|\boldsymbol{H}\|^2]$ is the channel fading power, and $\mathbb{E}[\text{Tr}(\boldsymbol{NN}^*)]$ is the noise power. The operators are denoted $\mathbb{E}(.)$ for expectation, $\text{Tr}(.)$ for trace and $(.)^*$ for complex conjugate. For a SISO scheme, (3.6) is reduced to yield

$$\gamma_s = \frac{\sigma_h \sigma_x}{\sigma_n^2} \tag{3.7}$$

and the dB expression of which is simply calculated as

$$\gamma_s = 10log_{10}\gamma_s \text{ (dB)} \tag{3.8}$$

The above symbol SNR is used to calculate the symbol probability of error P_{es}, while the bit SNR γ_b is typical in computing the bit probability of error P_{eb}. Assume that Gray coding is used for a reasonable extent of SNR values, the symbol power is divided equally among all bits so that one symbol error corresponds to exactly one bit error. Such assumptions are appropriate for M-ary signalling and lead to the following approximations

$$\sigma_b^2 = \frac{\sigma_x^2}{log_2 M} \qquad \gamma_b = \frac{\gamma_s}{log_2 M} \tag{3.9}$$

The spectral bit rate, or bandwidth efficiency, parameter R_s is now defined to learn how many bits per second R_b can be transmitted in a given bandwidth B [39, 42, 107]

$$R_s = R_b/B \text{ (bps/Hz)} \tag{3.10}$$

Larger R_s signifies a system with more bandwidth efficiency since it can transmit at a higher bit rate in each hertz of bandwidth. It is thereby evident that a good system is the one that at a given γ_s provides the highest R_s , or at a given R_s requires the least γ_s.

Continuing the above discussion, the probability of error is pronounced to play a major role, among other parameters, in the design and evaluation of wireless systems, and based on which the power efficiency can be deduced as well. The multifaceted relationship between SNRs ($\gamma_b$or γ_s) and error probability is phenomenally organic. All of M-ary systems basically tend to increase spectral bit rate at the expense of error probability or SNR. The power efficiency criterion of a particular signalling scheme generally specifies the SNR per bit that is essential to attain a certain error probability. The error probability that is commonly adopted for comparison of various signalling schemes is 10^{-4}, or sometimes 10^{-5}. The SNR required by a particular signalling scheme to achieve the above predefined error probability is a power efficiency criterion of that scheme. Other signalling schemes requiring lower SNR to achieve this error probability are therefore said to have a better PE [39, 107]. The power efficiency parameter can be recognised as the argument multiplied by either γ_b or γ_s in error probability expressions P_{eb} or P_{es}, respectively [42]. The above performance design criteria for common digital signalling schemes, such as binary PSK (BPSK), QPSK, MPSK and MQAM, and assuming coherent detection and AWGN environment are summarised in Table 3.2 below [39, 42, 107]. The right tail cumulative distribution function is given by $\mathrm{Q}(u) = \frac{1}{2}erfc(\frac{u}{\sqrt{2}})$ or $\mathrm{Q}(u) = \frac{1}{\sqrt{2\pi}}\int_u^\infty \mathrm{e}^{-x^2/2}dx$.

Table 3.2 Summary of approximate design performance of digital modulation schemes.

Modulation	R_s	**Power Efficiency**	P_{eb}	P_{es}
BPSK	1	1	$\mathrm{Q}\left(\sqrt{2\gamma_b}\right)$	-
QPSK, 4-QAM	2	1	$\mathrm{Q}\left(\sqrt{2\gamma_b}\right)$	$2\mathrm{Q}\left(\sqrt{\gamma_s}\right)$
MPSK	$log_2 M$	$sin^2\frac{\pi}{M}log_2 M$	$\frac{2}{log_2 M}\mathrm{Q}\left(\sqrt{2\gamma_b log_2 M}sin\frac{\pi}{M}\right)$	$2\mathrm{Q}\left(\sqrt{2\gamma_s}sin\frac{\pi}{M}\right)$
MQAM	$log_2 M$	$\frac{3log_2 M}{2(M-1)}$	$\frac{4}{log_2 M}\mathrm{Q}\left(\sqrt{\frac{3\gamma_b log_2 M}{M-1}}\right)$	$4\mathrm{Q}\left(\sqrt{\frac{3\gamma_s}{M-1}}\right)$

The experience of MIMO and fading channel situations can be tailored to the above approximations by embedding the expressions of (3.6) and (3.7) as relevant.

An illustrative simulation exercise is now conducted to assess the probability of bit error P_{eb}, which is also commonly called BER, with respect to M-ary signalling and

SNR variations. The size and type of M-ary signalling schemes are arbitrarily selected for a tandem group of 8, 16 and 32 PSK and QAM constellations. The values of P_{eb} are computed as per the entries of Table 3.2. The simulation settings also consider a coherent type detection of the form of ML algorithm, SISO configuration and AWGN contamination. The performance trends of theoretical P_{eb} are exhibited in Fig. 3.3.

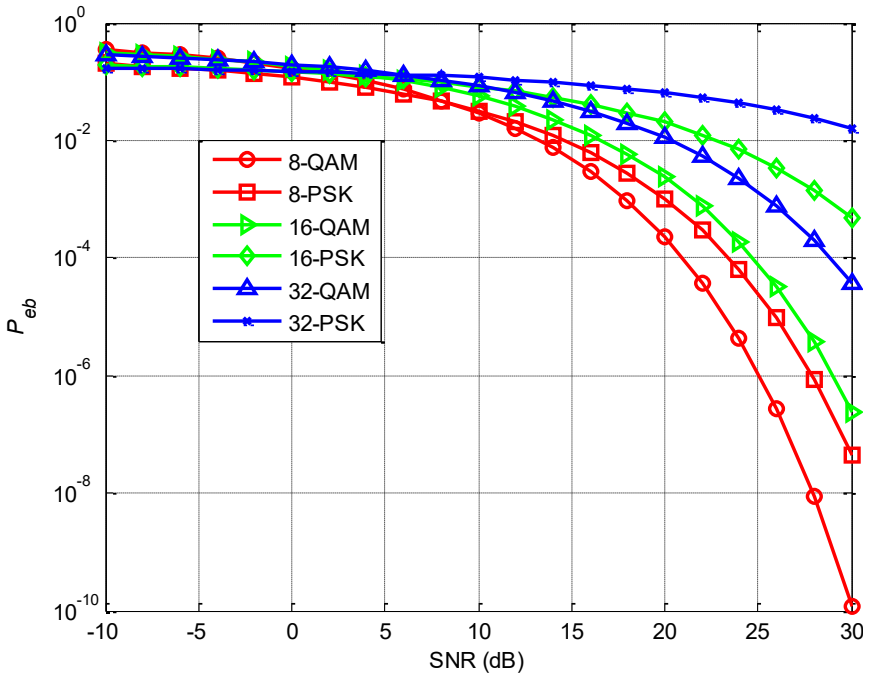


Fig. 3.3 Probability of bit error for arbitrary digital modulations using SISO in AWGN.

The first observation can be deduced from the above figure is that signalling schemes of larger size are prone to more errors than smaller size schemes for the same SNR values. Such behaviour is very well expected and mainly attributed to the dramatic changes in the minimum distances necessary for detection thresholds between constellation points. In as much as M increases, additional points assigned to each symbol which decreases the distances between adjacent constellation points. This eventually makes comparing distances based on predefined thresholds a tedious endeavour and hence more errors are generated at the output of the detection stage.

The second observation that can be drawn from Fig. 3.3 is related to the advantageous detection performance of MQAM compared to MPSK. It can be seen that P_{eb} for MQAM outperforms that of MPSK for the same M-ary sizes and SNRs. This can be easily interpreted in terms of the P_{eb} expressions given in Table 3.2 and identify

the PE argument associated to each of the signalling schemes. This argument is well perceived in the divisor $(M-1)$ of dominant $\mathrm{Q}(.)$ function per se. The dividend is inversely proportional to its governing function and hence decreases its value due to M-ary increase, which represents a bonus for MQAM since the $\mathrm{Q}(.)$ values decrease accordingly. Therefore, an increase in the SNR by certain amounts is entailed for MPSK to maintain the same performance as that of MQAM, and Table 3.3 below gives an overview of some of these penalty factors [107].

Table 3.3 SNR advantages of MQAM over MPSK.

M	$10log_{10}\left\{\left(\frac{3}{M-1}\right)\Big/sin^2\left(\frac{\pi}{M}\right)\right\}$
8	1.65
16	4.20
32	7.02
64	9.95

It can hence be concluded that MQAM has better power efficiency performance compared to MPSK of the same M-ary size and eventually is more privileged in practical reality.

3.3 Diversity and Combining Techniques

It has been shown earlier in the context that signals suffer from great degradation when transmitted through fading channels, as their original characteristics are almost completely deformed to the extent that they are not easily recognizable or detected. Therefore, diversity techniques are one way among others sought to combat such adverse situations and improve the reliability of communication without increasing either the transmitted power or bandwidth. They are mainly intended to alleviate the deleterious multipath fading effects and enhance transmission parameters like power, reliability, throughput capacity and other performance criteria [39, 107, 108]. These techniques are vital and have long been extensively used in wireless communication systems and their legacy is expected to last for many years to come.

3.3.1 Diversity System

The term “diversity system” generally refers to a system in which two or more closely related copies of some wanted signal are available. In situations where the desired signal suffers from deep fading, some kind of diversity reception can be employed to reduce the fade extent, in both amplitude and duration, and hence improve the communication quality. It is known that if two or more radio channels are sufficiently separated in space, frequency, or time, and sometimes in polarisation, then the fading on the various channels is more or less independent; i.e., it is then relatively rare for all the channels to fade together. When there are constraints on the transmitted power and bandwidth, then diversity reception is one way to improve the reliability of communication without breaching these constraints.

The following is a brief description of the main diversity techniques [39, 107, 108].

- Time diversity: involves transmitting the same signal two or more distinct times with probably some sort of delay between replicas. Such a simple form of time diversity is considered complex, computational, and deteriorates transmission throughput, despite the use of one antenna and one carrier.

- Frequency diversity: involves transmitting the same signal on two or more carrier frequencies. If these are sufficiently separated, the fading on the various signals is approximately independent. This method is economical in terms of antennas and infrastructure, but is bandwidth expensive.

- Space diversity: a single or multiple transmitting antennas furnish a signal to several well-separated remote receiving antennas; this method is also called antenna diversity. An example of two separated transmitting antennas, one of which transmits vertically polarised radiation and the other transmits horizontally polarised radiation. A single receiving reflector with two feed horns or dipoles is used to separate the vertical and horizontal received signals. By combining these two methods, full duplex space diversity system that requires only two reflectors at each end is achieved. In one form or another, space diversity has been the most commonly used technique of diversity communication and hence is explored further in what follows.

3.3.2 Spatial Combining

A simple representation of the linear space diversity combiner scheme is depicted in Fig. 3.4. Several space diversity combining and switching techniques are known in the literature and can be practically classified under four main categories, namely; the selection combining (SC), switch-and-stay combining (SSC) (sometimes also called threshold combining), equal-gain combining (EGC), and maximal-ratio combining (MRC) [39, 107-109]. As far as the combining methods are concerned, it is irrelevant in principle whether combining takes place before or after demodulation when the demodulation process is linear, but of vital concern, for example, in systems employing frequency modulation, where the demodulation process is inherently nonlinear. In this case, pre-detection combining yields a higher baseband SNR than post-detection combining, and hence the former is of special interest in the world of mobile radio systems since they mostly use frequency modulation.

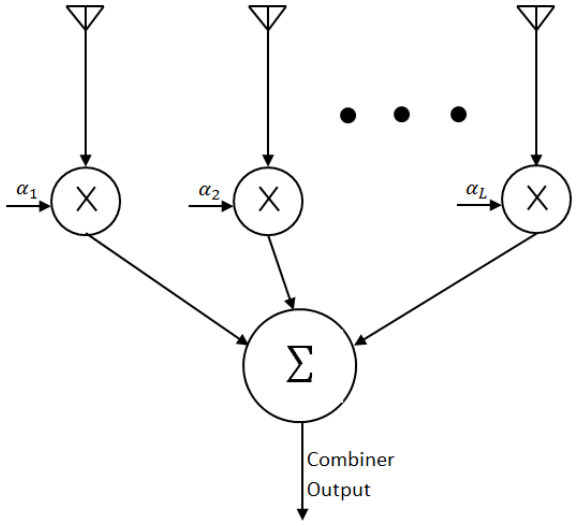


Fig. 3.4 Linear space diversity combiner.

The following briefly outlines the dominant techniques of space diversity combining.

- SC: The combiner selects the signal at the branch that experiences the highest SNR. In systems where one branch is active at a time, one receiver is needed to switch to the highest SNR branch. On the other hand, in systems that transmit continuously on different branches, a receiver at each branch is needed to monitor the SNR on each of them, so the output SNR is equal to that of the best incoming

signal. It is apparent that one receiver is needed for each branch and hence this would be a computationally demanding and time consuming process.

- SSC: This is a modified version of SC method, as one receiver is used to scan all branches sequentially and select the signal with SNR higher than a predefined threshold. If the SNR drops below the threshold on the selected branch, the combiner switches to another branch either randomly or systematically.

- MRC: This is also an improvement on SC and SSC in the sense that the output selection is performed by a weighing process. In MRC, the output is a weighted sum of all signal branches. The branch with the highest SNR will be assigned with a higher weight. The MRC can be used for coherent communication systems, since the weighting elements can be adjusted based on the CSI.

- EGC: This is a simpler form of MRC, where all signal branches are weighted equally and then added together. This method along with SC and SSC can be used for incoherent communication systems since the selection mechanism is not necessarily relying on the CSI.

The open literature suggests that the MRC optimality hinges on knowing the weighing factors of all branches and their values adaptively adjusted by probing into the attributes of the incoming signal. This materialises a perfect combining for maximum capacity over other combining techniques, as they cannot compete with the MRC performance, but if such awareness vanishes, then the MRC declines and acts just like any other combiner.

Two exercises have been worked out to demonstrate the benefits of space diversity techniques. The theoretical entries of Tables 3.4 and 3.2 are used to obtain the total SNR variations against different receive antenna settings (or transmit antennas as well as can be seen later in the context) and the P_{eb} with respect to SNR and receive antenna variations, as depicted separately in Figs. 3.5 and 3.6, respectively. While the general trends of Fig. 3.5 illustrates a continuous increase in the total SNR with respect to the number of antennas for all space diversity techniques, however, the MRC and EGC possess significantly better SNR gains than that of SC. Also, the fact that MRC outperforms EGC is very obvious and this is the main reason behind its practical preference, and also throughout the work of this study. As for the error detection, the

P_{eb} trends of Fig. 3.6 below clearly show that performance improves as the number of antennas increases using BPSK signalling in MRC and AWGN settings.

Table 3.4 Total SNRs for different space diversity combining schemes.

Type	γ_{Ts}	Notes
SC	$\bar{\gamma}_s \sum_{i=1}^{N_r} \frac{1}{i}$	Average SNR for all branches are the same $\bar{\gamma}_s$, co-phasing not required
EGC	$\frac{1}{N_r \sigma_n^2} \left(\sum_{i=1}^{N_r} \sigma_{s_i} \right)^2$	Equal noise for each branch, co-phasing as necessary
MRC	$\sum_{i=1}^{N_r} \gamma_{s_i}$	Total noise is considered, co-phasing required
SSC	$\gamma_s >$ Threshold	Co-phasing not required

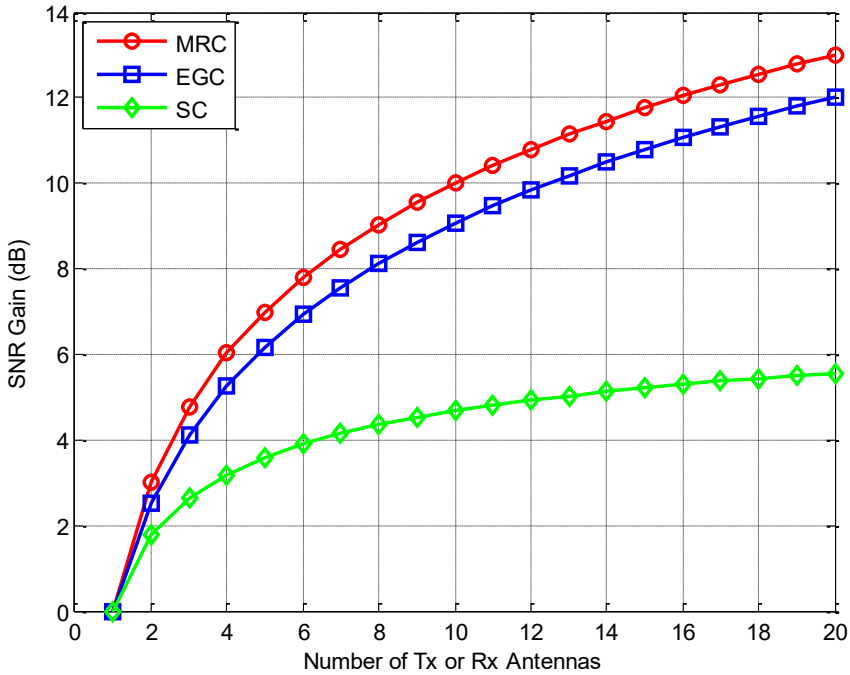


Fig. 3.5 SNR gains against number of receive, or transmit, antenna variations.

Before leaving this section, it is worth to learn about transmit diversity, which works almost similarly to receive diversity, and in which the transmit power is divided among multiple antennas. Such a technique is desirable where more space, power, and

processing capability is attainable on the transmitter side than on the receiver side, as in mobile and portable systems. The complexity and type of transmit diversity is highly coherent to the knowledge of complex channel gain on the transmitter side. If such channel knowledge is available, the system is quite similar to the receiver diversity [39, 110]. However, to maintain transmit diversity gain without having this channel information may require a combination of space and time coding techniques, which are explored in the next context. Inspiring the instrumental merits of the MRC technique, an insightful treatment of the concept and principles for maximum ratio transmission (MRT) was approached in [110].

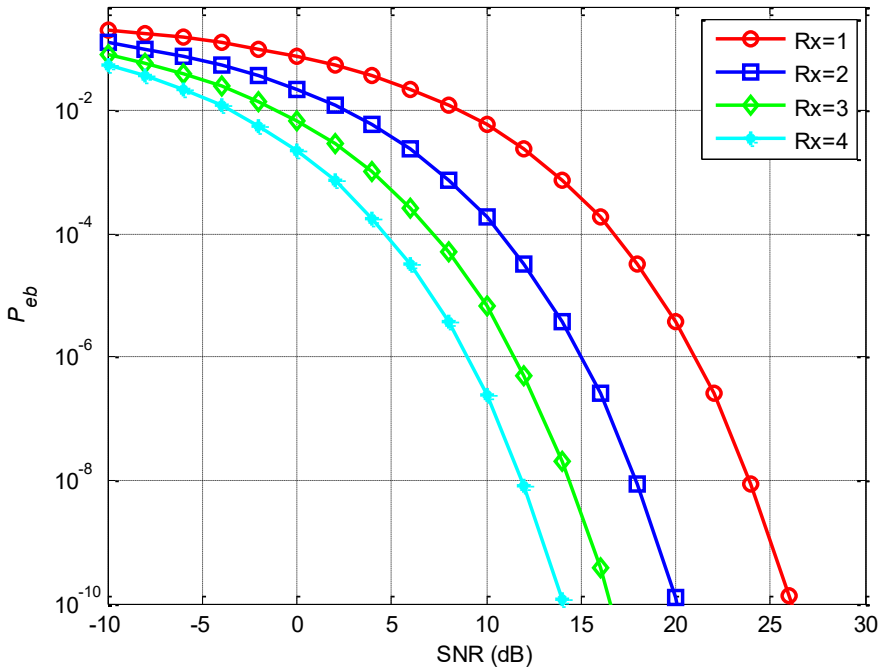


Fig. 3.6 Probability of BPSK bit error against SNR and receive antenna variations in AWGN.

3.4 Multiple-Antenna Technology

The demand for highly reliable and vast throughput in modern wireless communication systems quickly took advantage of the compelling features of various space diversity schemes. This paved the way for new advancements in the wireless world that heavily rely on exploiting multiple antennas on both transmit and receive sides. Collimated studies on the concentration of multiple-antenna configurations in advanced wireless communication systems have thrived accordingly [111]. The boost in

several transmission parameters rendered the multiple-antenna as a standardised industrial technology in almost all wireless systems. The amalgamation of multiple-antennas with multiple carrier systems, like OFDM, can furtherly achieve unbeatable advantages with respect to severely deteriorated fading conditions [46, 112, 113]. Distinctive spatiotemporal coding and combining techniques proved to be very useful in such brilliant amalgamations. Owing to their rich capabilities, whether united or solitary, the MIMO-OFDM technologies have become integral parts of a broad range of wireless communication systems, and their future is also expected to thrive in the emerging CR applications [10, 12, 13, 15, 16, 31, 55, 59-62, 91-94, 100].

3.4.1 MIMO

The sophisticated MIMO technology is spreading rapidly on all wireless fronts. Although such space diversity technique can take different forms, however, its central objectives are to deliver higher SNR, spectral, bandwidth and rate efficiencies [111-113]. A breadth of MIMO arrangements can readily be extended to the development of effective CR system and CRN settings. Further to the above, fresh studies suggested that MIMO schemes can be smartly reconfigured to append different modes of CR operation, such as concurrent sensing and transmission subject to certain conditions being met [114]. Combined with particular types of network coding, they also can assist in elevating the effectiveness of secured transmission of data packets over CRNs [115].

Nonetheless their long existence, MIMO has only become popularly fascinating and actively explored in the last decades, precisely since the end of 1990s [60, 111]. The growing interest was mainly fuelled by the pioneering works of STBC, which was the spark that ignited enormous research efforts on this vibrant technology. Surprisingly, despite being intensely researched for relatively a short period; MIMO has already been adopted in many wireless communication systems' standards, such as 3rd-generation (3G), 3rd Generation Partnership Project LTE (3GPP-LTE) and 4th-generation (4G). As the new concept of massive MIMO is gradually shaping up, it has been recently admitted to the 5th-generation (5G) to push the operational boundaries beyond imagination. Many IEEE 802.x industrial standards have thereby grossly surrendered their core air interfaces to the MIMO transmission to earn the plenty bonuses out of it [111].

MIMO are space-time wireless techniques that can be used to gain either full diversity or full rate (capacity), or both [39, 107, 111, 113, 114]. In a dense multipath scattering environment, a MIMO system features the diversity of multiple antennas separated in space. MIMO systems may come in different forms to gain either a full diversity to combat signal fading or to gain a full rate to increase throughput. They can hence be divided into three main groups. The first group aims to improve the power efficiency by maximising space diversity. Such techniques include time diversity, STBC and space-time trellis codes (STTC). The second group aims to increase the coding gain by using a layered approach, the most popular example of which is called Vertical-Bell Labs Layered Space-Time (V-BLAST), where full diversity is usually not achieved. The third and last group exploits the CSI at the transmitter. It splits the channel into parallel paths using SVD and uses the decomposed unitary matrices as pre and post processing at the transmitter and the receiver to achieve near capacity.

Bearing all the above in mind, let now delve into some basic analysis to grasp the main viewpoints. A wireless communication system employing N_t transmitting antennas and N_r receiving antennas, like the one shown in Fig. 3.1 and mathematically expressed in 3.3 and 3.4, is generally called a MIMO system, and the resulting spatial channel in such a system is called MIMO channel [39, 107]. Different special versions can be derived from a MIMO system. The first is constituted of $N_t = N_r = 1$ and called a SISO system and its associated channel is of SISO type. The second comprises $N_t = 1$ and $N_r \geq 2$ and the system and channel of which called SIMO in this case. The third and last case consists of $N_t \geq 2$ and $N_r = 1$ and the resulting system and corresponding channel are of MISO type. Also it is worth to note that all combining techniques can be applied on both the transmitter and the receiver sides.

To evaluate the MIMO capacity (achievable rate), the treatment below is limited to frequency-nonselective channels only. For the sake of simplicity, the CSI is assumed known to both the transmitter and receiver, which is commonly denoted by CSIT and CSIR, respectively [39, 107]. Firstly, for the AWGN channel, the capacity is governed by the famous Shannon's formula, where the unit (bps) is bit per second [2, 12, 39, 107]

$$\mathcal{C}_{AWGN} = Blog_2(1 + \gamma_s) \quad \text{(bps)} \tag{3.11}$$

Shannon's theorem stated; there exists a code that achieves data rates arbitrarily close to capacity with arbitrarily small probability of error. Conversely, any code with rate

$\mathcal{R} > \mathcal{C}$ has a probability of error bounded away from zero [39, 107]. For static SIMO, MISO and MIMO channels, the capacities are defined below, respectively [12, 39, 44, 55, 60, 107]

$$\mathcal{C}_{SIMO} = Blog_2\left(1 + \gamma_s \sum_{i=1}^{N_r} |h_{i1}|^2\right) \quad \text{(bps)} \tag{3.12}$$

$$\mathcal{C}_{MISO} = Blog_2\left(1 + \gamma_s \sum_{i=1}^{N_t} |h_{1i}|^2\right) \quad \text{(bps)} \tag{3.13}$$

$$\mathcal{C}_{MIMO} = Blog_2 det\left(\boldsymbol{I}_{N_r} + \boldsymbol{H}\boldsymbol{R}_{\boldsymbol{x}}\boldsymbol{H}^{\mathrm{H}}\right) \quad \text{(bps)} \tag{3.14}$$

where $\boldsymbol{R}_{\boldsymbol{x}}$ is the covariance of input sequence, det(.) denotes determinant and $(.)^{\mathrm{H}}$ is the Hermitian operator. For *iid* signals among transmitters with equal symbols' power, the $\boldsymbol{R}_{\boldsymbol{x}}$ is given by $\left(\gamma_s \boldsymbol{I}_{N_t}/N_t\right)$. The MIMO capacity (3.14) is hence revised as below

$$\mathcal{C}_{MIMO} = Blog_2 det\left(\boldsymbol{I}_{N_r} + {}^{\gamma_s}/_{N_t}\,\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right) \quad \text{(bps)} \tag{3.15}$$

The theoretical computation results of these capacities for arbitrary multi-antenna configurations with respect to γ_s variations are depicted in Fig. 3.6 below. These runs assumed B equals 1, or normalised capacity, and the channel coefficients are *iid* zero-mean and complex-valued with unit variance.

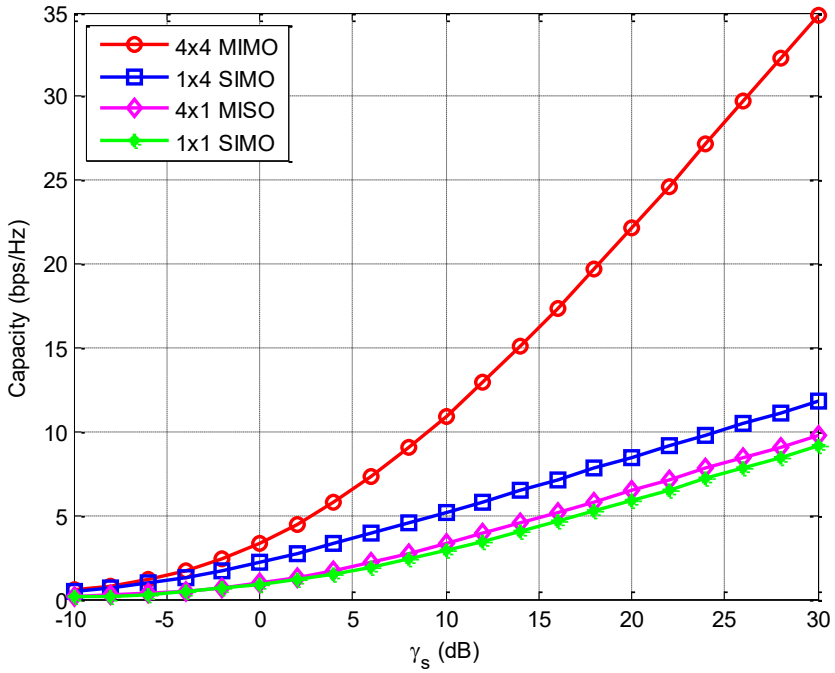


Fig. 3.7 Normalised capacity of multi-antenna systems with respect to SNR.

The capacities' trends in Fig. 3.7 evidently validate the MIMO superiority compared to other schemes of the same transmit or receive antenna counts. SIMO also performs better than MISO for the same number of antennas. This is due to the power split evenly among transmit antennas in MISO, while the power is collected by all receive antennas in SIMO.

3.4.2 STBC

The STBC was devised as a promise for more efficient and reliable MIMO systems. The major discovery of STBC was attributed to the leading work of Alamouti and Tarokh, among a few other separate and parallel efforts [39, 107, 111]. The simplest STBC scheme using 2×1 MISO was first proposed by Alamouti [39, 107, 112, 113, 116], and later was relaxed to higher-order STBC using more arbitrary antennas in the context of general orthogonal designs [107, 117-119]. Further readings on this vital topic can be found in two fine textbooks [118, 119], besides the general approach of [39, 107]. A simplified diagram of Alamouti diversity combining is shown in Fig. 3.8.

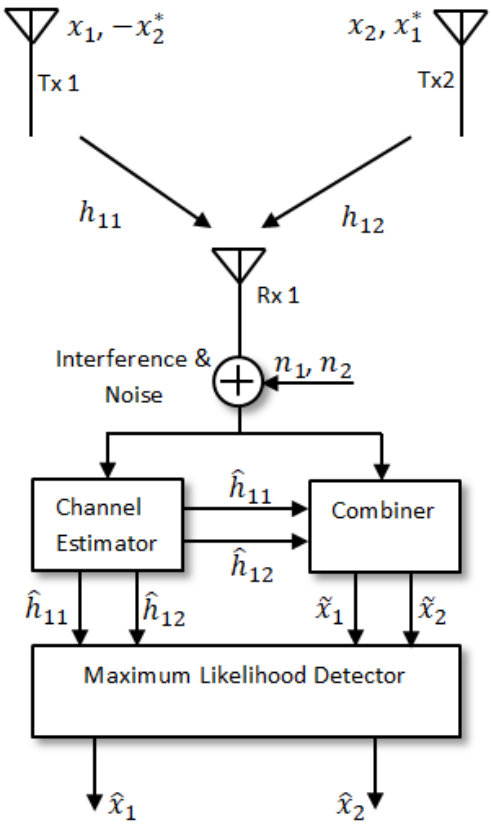


Fig. 3.8 Alamouti STBC scheme using 2×1 MISO arrangement.

Alamouti proposed two transmit antennas and one receive antenna to achieve the same diversity order as can be provided by MRC with one transmit antenna, and two receive antennas. Alamouti also demonstrated that the proposed scheme can be generalised to two transmit antennas and N_r receive antennas to provide a diversity order of $2N_r$ [116]. The proposed scheme was claimed not to require any bandwidth expansion or any feedback from the receiver to the transmitter and its computation complexity is similar to MRC. Encoding is performed in both time domain and space domain. That is two consecutive time periods will be used to transmit signals on two transmission antennas. These signal sequences are assumed to be orthogonal and uncorrelated to each other and hence their inner product is equal to zero. At a given symbol period, two signals are simultaneously transmitted from the two antennas. The signal transmitted from antenna one is denoted by x_1 and from antenna two by x_2. During the next symbol period signal $-x_2^*$ is transmitted from antenna one, and signal x_1^* is transmitted from antenna two, where $(.)^*$ means conjugate operator. This can be expressed in terms of the code generator and channel matrices [39, 107, 116, 118, 119]

$$\boldsymbol{G} = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}, \qquad \boldsymbol{H} = [h_{11} \quad h_{12}] \tag{3.16}$$

In general, the STBC code generator matrix $\mathbf{G}$ can be expressed linearly as follows

$$\boldsymbol{G} = \sum_{i=1}^{M} \Re(x_i)\widetilde{\boldsymbol{A}}_i + j\Im(x_i)\widetilde{\boldsymbol{B}}_i \tag{3.17}$$

or, equivalently substituting $x_i = \Re(x_i) + j\Im(x_i)$ yields

$$\boldsymbol{G} = \sum_{i=1}^{M} x_i \boldsymbol{A}_i + x_i^* \boldsymbol{B}_i \tag{3.18}$$

where a set of M complex symbols (MPSK or MQAM) is mapped onto a code matrix $\{x_1, x_2, \dots, x_M\} \rightarrow \boldsymbol{G}$ using arbitrary basis matrices $\{\boldsymbol{A}_i, \boldsymbol{B}_i\}$, with $\boldsymbol{A}_i = \left(\widetilde{\boldsymbol{A}}_i + \widetilde{\boldsymbol{B}}_i\right)/2$ and $\boldsymbol{B}_\mathrm{i} = \left(\widetilde{\boldsymbol{A}}_i - \widetilde{\boldsymbol{B}}_i\right)/2$ and all these matrices are specific code design dependent and have dimension $N_r \times N_t$ of real entries. Now, following the same approach given in [118, 119], (3.3) can be an explicit linear function in the data symbols x_i. Such alternation reduces the decoding complexity due to the rich structure of $\boldsymbol{G}$ and hence the

search needs not to be applied over all M symbols. The following vectors can be defined

$$\begin{aligned} \Re(\boldsymbol{X}) &= [\Re(x_1) \dots \Re(x_M)]^{\mathrm{T}} \\ \Im(\boldsymbol{X}) &= [\Im(x_1) \ \dots \ \Im(x)]^{\mathrm{T}} \\ \boldsymbol{X} &= [\Re(\boldsymbol{X}) \quad \Im(\boldsymbol{X})]^{\mathrm{T}} \end{aligned} \tag{3.19}$$

and

$$\begin{aligned} \boldsymbol{F}_a &= [\mathrm{vec}(\boldsymbol{H}\boldsymbol{A}_1) \quad \dots \quad \mathrm{vec}(\boldsymbol{H}\boldsymbol{A}_M)] \\ \boldsymbol{F}_b &= [j(\mathrm{vec}(\boldsymbol{H}\boldsymbol{B}_1)) \quad \dots \quad j(\mathrm{vec}(\boldsymbol{H}\boldsymbol{B}_M))] \\ \boldsymbol{F} &= [\boldsymbol{F}_a \quad \boldsymbol{F}_b] \end{aligned} \tag{3.20}$$

which after direct substitution will alter (3.3) to the following

$$\boldsymbol{Y} = \boldsymbol{F}\boldsymbol{G}^{\mathrm{T}} + \boldsymbol{N} \tag{3.21}$$

where $\Re(.)$ and $\Im(.)$ stands for real and imaginary, $(.)^{\mathrm{T}}$ transpose operator, $\mathrm{vec}(.)$ is a vector obtained by stacking the columns of the matrix argument on top of each other and $\boldsymbol{F}$ is the new channel matrix for STBC of code generator matrix $\boldsymbol{G}$, while $\boldsymbol{N}$ does not alter its statistical nature. The $\boldsymbol{F}$ matrix includes effects of both the channel and coding schemes. The above equivalent representation is credited to an important feature of many STBC algebraic constructions, which is the exchangeability of correlation structure at the transmitter to permit a comparably equivalent restructure at the receiver.

If M-ary distinct symbols are transmitted during one code block of K number of bit periods, then the STBC coding rate would be $\mathcal{R} = M/K$. An STBC is said having full rate if $M/K = min(N_t, N_r)$, where N_t and N_r are the numbers of transmit and receive antennas, respectively. Suppose a BPSK is aimed for transmission, the coding rate $\mathcal{R}$ equals constellation size/STBC time interval, which is $M/(K = M) = 2/2 = 1$ and hence data full rate is achieved in addition to spatial full diversity order. This is the main virtue of Alamouti code.

To decode the signal model of (3.21) on the receiver side, a space-time matched filter (STMF) $\boldsymbol{F}^*$ is typically utilised alongside the MRC to enhance the signal strength while admitted to a decoder. The most commonly utilised decoder is of the type of

MLE, which extracts all available spatiotemporal diversity order in the channel. In omitting the time for easiness; the received signal in this case is expressed by

$$y = \boldsymbol{F}^{*}\boldsymbol{Y} = \|\boldsymbol{H}\|_{\mathrm{F}}^{2} x_i + n_i, \quad \forall i \in [1, M] \tag{3.22}$$

where $\|.\|_{\mathrm{F}}$ denotes Frobenius norm.

Owing to the properties of $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$, to have the squared Frobenius norm of $\boldsymbol{H}$ given by $\|\boldsymbol{H}\|_{\mathrm{F}}^{2} = \mathrm{Tr}(\boldsymbol{H}\boldsymbol{H}^{*}) = \sum_{i=1}^{N_r}\sum_{j=1}^{N_t}|h_{ij}|^{2}$, the signal representation in (3.22) can be easily verified. It can hence be recognised that the above linear processing has rendered the STBC model to be equivalent to a set of M independent parallel transmissions over flat channels with coefficients $\|\boldsymbol{H}\|_{\mathrm{F}}^{2}$. The extraction of x_i from $\boldsymbol{Y}$ is simply attended by decoding of x_i at the output of MLE stage for $i \in [1, M]$. Such simple linear MLE processing and decoding results in a reduction of the computational load from M^2 metrics to $2M$ metrics [107], which is yet another virtue of Alamouti code. To compute the probability metrics, the SNR entries of Table 3.2 are merely multiplied by the factor $\|\boldsymbol{H}\|_{\mathrm{F}}^{2}$.

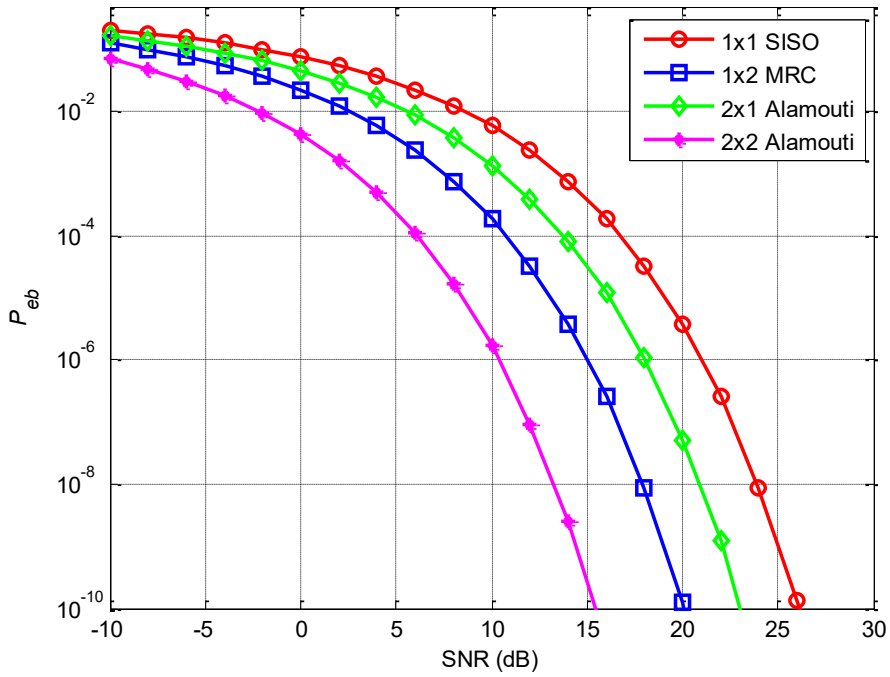


Fig. 3.9 BPSK probability of bit error for Alamouti and MRC against SNR in AWGN channel.

The P_{eb} theoretical metrics for SISO, Alamouti and MRC schemes are computed with respect to SNR and using BPSK signalling transmitted in a flat-fading channel and immersed in AWGN as revealed in Fig. 3.9. The perception of Fig. 3.6 above can also be drawn here with respect to the obvious space diversity gain of both Alamouti and MRC compared to SISO, however, the Alamouti code is slightly worse than MRC by approximately 3 dB penalty in performance. Despite this downside, which solely occurs in situations of having the same number of multiple antennas on either transmit side or receive side, however, as the antennas collectively increase on both sides, the upside improvement so quickly picks up to much better levels as unveiled in the same Fig. 3.9 above for the 2×2 Almouti code.

It is important to recognise that the pioneering work of Alamouti STBC attracted a great attention in industry and academia due to its appealing feature of low decoding complexity. It is believed to be the corner stone for highly sophisticated STBC algorithms mainly aimed at full diversity in coherent and non-coherent wireless communication systems targeting higher data rates at low costs. However, it has been shown that a complex orthogonal design and the corresponding STBC, which provides full diversity and full rate transmission, are not possible for more than two antennas [111, 120]. Therefore, another coding structure using a quasi-orthogonal design was suggested to come up with a new code specified as quasi-STBC, or QSTBC [111, 118, 120]. Using such a code, pairs of transmitted symbols can be decoded separately and hence achieve higher transmission rates, however, at the expense of sacrificing the full diversity. Alamouti and other STBC are generally highly respected assets in modern MIMO systems.

3.4.3 SVD Channel Decomposition

The SVD, or eigenvalue decomposition (EVD), is an effective generalisation of the PCA in the statistical data analysis, or Karhunen-Loeve Transform (KLT) in the stochastic estimation and detection theories alongside the pattern recognition. It has a pair of unitary orthonormal matrices, instead of one as in PCA and KLT, and it is widely used for de-noising and channel parsing [1, 39, 46, 93, 106, 107, 113, 118, 119].

1) MIMO Decomposition

The $N_r \times N_t$ channel matrix $\boldsymbol{H}$, with rank $\mathcal{R} = min(N_t, N_r)$ as defined earlier in the context, has an SVD representation given by $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^*$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $N_r \times \mathcal{R}$ left and $N_t \times \mathcal{R}$ right matrices to satisfy $\boldsymbol{U}\boldsymbol{U}^* = \boldsymbol{V}\boldsymbol{V}^* = \boldsymbol{I}_{\mathcal{R}}$, $\boldsymbol{\Sigma} = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_{\mathcal{R}}]$ with $\sigma_i \geq 0$ and $\sigma_i \geq \sigma_{i+1}$, where σ_i is the i^{th} singular-value of the channel. The columns of $\boldsymbol{V}$ and $\boldsymbol{U}$ are also known as the left input and right output singular vectors, respectively. Now, $\boldsymbol{H}\boldsymbol{H}^*$ is an $N_r \times N_r$ semi-definite Hermitian matrix. Let the eigen decomposition of $\boldsymbol{H}\boldsymbol{H}^*$ be $\boldsymbol{Q}\boldsymbol{\lambda}\boldsymbol{Q}^*$ where $\boldsymbol{Q}$ is an $N_r \times N_r$ matrix satisfying $\boldsymbol{Q}^*\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}_{N_r}$ and $\boldsymbol{\lambda} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_{\mathcal{R}}]$ with random variable eigenvalues $\lambda_i \geq 0$, sorted in a descending order $\lambda_i \geq \lambda_{i+1}$ and have Chi-square distribution and $2N_tN_r$ degrees of freedom [118, 119]. Then, $\lambda_i = \sigma_i^2$ for $i \in \{1, 2, \dots, \mathcal{R}\}$ and $\lambda_i = 0$ for $i \in \{\mathcal{R} + 1, \mathcal{R} + 2, \dots, N_r\}$.

Assume the channel is known and may be retrieved through linear processing at both ends. Hence, a precoding is carried out at the transmitter by multiplying the data sequence by $\boldsymbol{V}$ to generate $\widetilde{\boldsymbol{X}}$, while a post coding is applied on the receiver by multiplying the data sequence and noise by $\boldsymbol{U}^*$ to recover the original data sequence and modified noise variable $\widetilde{\boldsymbol{N}}$ without affecting its statistics. This alters (3.3) to have the following form

$$\boldsymbol{Y} = \boldsymbol{\Sigma}\widetilde{\boldsymbol{X}} + \widetilde{\boldsymbol{N}} \tag{3.23}$$

This formula can be revised to envisage the channel knowledge on transmitter and hence $\boldsymbol{H}$ can be explicitly collapsed into $\mathcal{R}$ parallel SISO channels. Using a proper STMF of the type $\boldsymbol{\Sigma}^*$, or MRC, the SNR can be maximised and hence renders the following [118, 119]

$$y = \sum_{i=1}^{\mathcal{R}} \sqrt{\lambda_i}\, x + n \tag{3.24}$$

which means a MIMO receiver can see one complex symbol at a time.

2) STBC Decomposition

Applying the SVD on the STBC formation is straightforward. The metric $\|\boldsymbol{H}\|_{\text{F}}^2 = \sum_{i=1}^{\mathcal{R}} \lambda_i$ can be thought of as the total power gain of the channel, yielding (3.22) to be

$$y = \sum_{i=1}^{\mathcal{R}} \sqrt{\lambda_i}\, x_j + n_j \;, \quad \forall j \in [1, M] \tag{3.25}$$

which means that all symbols x_j can see the same channel at one instant in time. Equation (3.25) shows that the STBC lends itself very easily and directly to the SVD channel parsing structure. Furthermore, it obviously has extra power advantage over an ordinary MIMO system due to the built-in coding process, as indicated earlier in the context. Such additional gain is attributed to the direct reliance on the M-ary constellation symbols observing the same channel eigenvector λ_i at a time. Engaging the STBC hence doubles the coding gain compared to that of a MIMO system, irrespective of diversity gain and channel fading conditions. This conforms to the coding gain of around 3 dB is readily achievable for 2×2 STBC [55, 119], which increases by $10log_{10}M$ for higher order codes. It is also worth to note that the STBC coding rate resembles the minimum antennas' geometry to achieve full diversity, which is given by $\mathcal{R}$. Therefore, it is conventionally known that most complicated STBC schemes, other than Alamouti, can achieve either full rate or full diversity.

Having said that, the number of eigenvalues for the Alamouti case is four times higher than that for the general MIMO system due to the STBC coding being performed over two time intervals, and the observation is divided into real and imaginary parts. For Alamouti code, the four largest eigenvalues are significantly larger than the others, while for the general MIMO case, it can be noted that two of the eigenvalues are significantly larger than the others, because the channel matrix has rank two [92]. In addition to that, the STBC practice is considered robust in case of poor channel reporting due to fading effects. Therefore, a receiver not necessarily always needs the channel knowledge to perform its duties and can work in a blind or semi-blind fashion when the STBC is involved [55].

3.5 Multiple-Carrier Technology

The multi-carrier technology was invented in the early 1960s as an effective tool to overcome the negative effects of channel fading [39, 107, 112, 113]. The main idea of the multi-carrier method is to divide the transmission bit stream into vast number of multiple sub-streams and send these over many different sub-channels separately.

3.5.1 OFDM

OFDM is widely accepted as the best candidate multi-carrier technology for its great flexibility and spectral efficiency and is currently used in many wireless communication systems. Recall that the main purpose of using OFDM modulation is to turn the frequency-selective channel into a set of parallel flat-fading channels and hence the inter-symbol interference (ISI) effect could be totally removed. This can be achieved by extending the OFDM symbol with a cyclic prefix to mitigate the channel time spread. In addition to that, OFDM enjoys further advantages including robustness against narrowband interference, scalability, and ease of implementation using fast Fourier transform (FFT). So much as that, the OFDM has become widely adopted in almost all air-interface standards focusing on CR applications [121-123]. Due to the sparkling features of both technologies, the merger of MIMO-OFDM promises the fulfilment of challenges set forward by xG demanding trend of very high speed and reliable wireless communication systems [124].

This subsection highlights the modelling of OFDM signal that is elementary to some parts of this study, and on no account to attempt the in-depth coverage of such an extensive topic. Consider an OFDM scheme where a block of K data symbols is transmitted in parallel on N_c subcarriers. The time duration of an OFDM symbol is K times larger than that of a single-carrier system. The cyclic prefix and suffix augmented guard samples can be added, but not considered in this brief analysis. The OFDM transmission system can be effectively realised in discrete time by employing the inverse fast Fourier transform (IFFT) to act as a modulator and the FFT to act as a demodulator. The transmitted data are the "frequency" domain coefficients, and the samples at the output of the IFFT stage are "time" domain samples of the transmitted sequence. Suppose there is a K-length data symbol block of the form $\boldsymbol{X} = \{X(1), X(2), \dots, X(K)\}$. Applying the IFFT on this data block generates the time domain sequence $\boldsymbol{x} = \{x(1), x(2), \dots, x(K)\}$ represented by [39, 107, 112, 113]

$$\{x(k)\} = \text{IFFT}_{N_c}\{X(k)\} \tag{3.26}$$

The input data sequence $x(k)$ is modulated by a particular M-ary constellation mapping process, such as MPSK or MQAM, resulting in the complex symbol stream $\boldsymbol{X}(k)$. It is important to note that IFFT and FFT must have an equal number of subcarriers and the

spacing between any two subcarriers is $1/N_cT_s$, which is significantly less than the time of one symbol T_s. Applying the FFT retrieves the original time sequence at the receiver side.

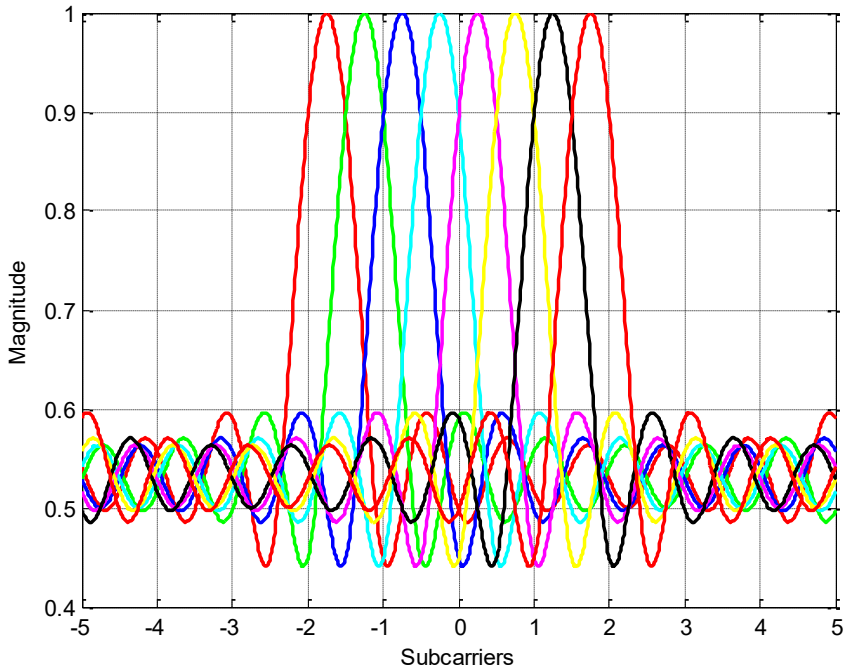


Fig. 3.10 Frequency spectrum of 8 orthogonal subcarriers of an OFDM transmit signal.

A simplified frequency response of 8-subcarrier OFDM modulation is synthesised as in Fig. 3.10. The orthogonal phenomenon is obvious by having a look at the peak magnitudes of particular subcarriers corresponding to nulls in the remaining subcarriers, and hence the ISI effects are eliminated thoroughly. The likelihood of having a great portion or all subcarriers experience fading at the same time is very rare and hence this intensifies the invulnerability of OFDM system against such detrimental situations.

3.6 Fading Channels

As described earlier in the context, a wireless communication system is intended to transfer digital messages via free space using various combinations of signalling and antenna schemes. But the fact is such free space is not available unless going far and beyond the atmosphere zone. Wireless transmissions at the liveable physical levels are normally impaired by many obstructions such as buildings, towers, small objects and

people themselves, which create signal fading, shadowing, echoing and reflections. Most challenging transmissions are encountered when either one or both of transmitter and receiver are in a continuous move, since the signal impairments in such scenarios will be vigorously changing and hence difficult to characterise fully. Therefore, most wireless communication systems suffer problems originating from adverse multipath fading in transmission channels that is more prone to dynamic variations. The severity of such fading is highly dependent on terminal mobility and obstructions [39, 45, 107, 118, 119].

Spurred by the aforesaid challenges, an appropriate channel model and simulation approach needs to be investigated to account for these channel effects. As explained earlier in the context, the SDR is envisioned as an enabling technology that can flexibly adapt against stringent channel conditions to maintain an adequate QoS. Vital CSI attributes can hence be utilised to allow the SDR to reconfigure its internal parameters and combat channel variations. The SDR within the CR framework can employ the ACM technique to cope with such kinetics. It is therefore imperative in the assessment of wireless communication systems to retain the design and analysis of channel fading as accurately as possible [42, 44].

Many recent studies are devoted to investigate such channel characterisation and simulation in wireless communication systems. A good survey on the FSMC origin of developments can be found in [125] and the references therein. It has been indicated that such statistical modelling can be traced to the initial efforts by Gilbert for a two-state channel crossover in wireline telephone circuits with burst-noise, and then improved by Elliot, in the early 1960s. This model is commonly known as the Gilbert-Elliot channel in the literature and it is intended to approximate the Rayleigh channel fading behaviour by only two states of channel quality. The basic idea behind such conceptual approximation is to find a tractable methodology for the formulation and calculation of channel information capacity and associated BER. Therefore, the FSMC approach can rather be called as an information-theoretic channel approximation.

Furthermore, in the Gilbert-Elliot channel model, each state represents a specific channel quality which is either noisy or noiseless. Generally, a binary symmetric channel (BSC) with a given crossover probability can be associated with each state so that the channel quality for each state can be identified. The Rayleigh fading effect in

this case is assumed to be wide stationary, which means the crossover and transition probabilities have fixed values not altered by time, and the FSMC is of the 1st-order type. The Gilbert-Elliot channel model has been widely espoused for performance studies in wireless fading channel environments. However, the Gilbert-Elliot channel has severe limitations especially in cases where channel characteristics are highly likely to change dramatically [126].

Several studies on the FSMC channel modelling can be identified in the literature, but only a few are described here. An attempt to establish a connection between Rayleigh fading channels and their FSMC counterparts can be found in [126]. In this study, the SNR was partitioned into a finite number of intervals, corresponding to the FSMC model. The zone between any two levels represents the fading channel state and hence the transition and crossover probabilities from one state to another can be interrogated using such an analytical-approximation approach. A policy to partition the received SNR into a finite number of states as per the time duration of each state was developed and analysed [127].

Alternatively and instead of the SNR range, the dynamic range of channel fade amplitude was considered for partitioning in [128]. It expanded the FSMC by introducing intermediate channel states in between adjacent symbol epochs of the actual de-interleaver output. Such expansion claims that the FSMC is usable for real situations of non-interleaved channels with fast fading (fast Doppler frequency to symbol-transmission rate ratio), or interleaved (correlated) channels, such as in diversity combining. The validity of adopting the 1st-order FSMC was examined using an alternate approach of autocorrelation function over consecutive data samples [129]. A different approach using adjacent transition was proposed to construct an FSMC model to represent the Rayleigh fading channel [130]. The adjacent transition method generally differs from the equal probability in [126] and equal duration in [127].

From many arbitrary examples on the FSMC viability, two are selected for exposure here. The first is on the FSMC modelling for wireless transmission losses that can be implemented to discriminate between wireless and congestion related losses in data networks using real channel traffic [131], while the second is on the FSMC-SS policies in CR systems. These systems leverage past sensing outcomes of several cooperating SUs to decide which channel of PUs should be sensed by each SU at a given time [132].

In recognition of the distinctive features of the FSMC practices, such cornerstone applications represent the new research trends in wireless communication systems.

3.6.1 Fading Channel Modelling

Time-varying multipath channels can be either time or frequency selective, or both together and as such are labelled as doubly selective. The speed of time and frequency variations determines whether a channel is of a slow fading or fast fading nature [39, 45, 107, 118, 119]. Generally, commercial land wireless communication systems are assigned predefined channels with small frequency bandwidths. Hence, our approach will consider a narrowband baseband signal over a flat-fading Rayleigh channel only.

Consider the linear dynamic model for a received baseband signal at the output of MF described as given in (3.1) above, but now the time index is reserved to reflect on channel changes. As before, $h(k)$ is the fading channel envelope, which is a recursive complex-valued random process, and $n(k)$ is the AWGN $\sim\mathcal{CN}(0,\sigma_n^2)$. The amplitude a and phase θ components of the fading channel are governed by $h(k) = a(k)e^{j\theta(k)}$. The fading channel can be decomposed into its real and quadrature components $h(k) = h_r(k) + jh_q(k)$ and each is given by $\sim\mathcal{N}(\mu,\sigma_h^2)$. The mean value μ represents the LOS propagation component; if it is zero, the fading process is called Rayleigh fading; otherwise it is called Rician fading. In multipath channels, the fading amplitude a has a Rayleigh PDF denoted by [39, 45, 125, 107, 133]

$$f_h(a) = \frac{a}{\sigma_h^2} exp(-\frac{a^2}{2\sigma_h^2}) \ , \ \ a \in [0,\infty) \tag{3.27}$$

and the fading channel phase has a uniform PDF as below

$$f_h(\theta) = \frac{1}{2\pi} \ \ , \ \ \theta \in [0,2\pi) \tag{3.28}$$

In the same manner, but with slightly different portrayal, such a Rayleigh PDF can also be implied for the fading SNR random process [127, 130, 131]. Even though multipath fading gains and their real and imaginary components are driven by uncorrelated WSS AWGN, it is known that the time variations of such channels interact in a correlated manner. Such inter-correlation will determine the channels' statistical,

time and frequency characteristics. It is hence essential to seek a tractable mathematical model to accurately describe the dynamic time variations of fading channels. The most widely accepted statistical model in such cases was developed by Jake and Clark [39, 107, 125, 129, 133]. In the Jake and Clark model, the autocorrelation function (ACF) of the real and imaginary components of a fading channel gain is given by

$$\begin{aligned} R_h(T_s) &= E\{h_r(k)h_r^*(k-T_s)\} \\ &= E\{h_q(k)h_q^*(k-T_s)\} \\ &= \sigma_h^2 J_o(2\pi f_D T_s) \end{aligned} \tag{3.29}$$

where $J_o(.)$ is the zeroth-order Bessel function of the first kind, $f_D = (v/c)f_o$ is the Doppler spread frequency, v is the terminal travelling speed, f_o is the carrier frequency, c is the speed of light, and T_s is the channel symbol duration. The PSD of the above ACF is denoted by

$$S(f) = \begin{cases} \dfrac{\sigma_h^2}{2\pi f_D} \dfrac{1}{\sqrt{1-\left(\dfrac{f}{f_D}\right)^2}} & \text{for } |f| < f_D \\ 0 & \text{otherwise} \end{cases} \tag{3.30}$$

The channel fading characteristics are determined by the Doppler frequency due to the motion of a mobile terminal. In slow or flat-fading channels, the channel coherence time $T_c \approx 1/f_D$ is larger than the symbol period, T_s, alternatively $f_D T_s \ll 1$. For extremely fast-fading channels $f_D T_s \gg 1$, however, such channels are not regularly encountered in reality.

3.6.2 AR Channel Modelling

The AR filters are finite impulse response (FIR) structures commonly utilised to approximate fading channels of particular time and frequency responses. This is due to the traceable computation of their parameters and correlation properties. Let us assume we have a frequency-nonselective fading channel with L resolvable paths. This channel is anticipated to be slowly varying and constant during the observation interval. By

using a conventional tapped delay line model with tap spacing equal to T_s, the fading formula for real and imaginary components is given by [1, 125, 133, 134]

$$h(k) = -\sum_{l=1}^{L} a_l h(k-l) + w(k) \tag{3.31}$$

where $\{a_1, a_2, \dots, a_L\}$ are the AR(L) filter coefficients of order L, $|a_l| < 1$, and $w(.)$ is a complex AWGN $\sim \mathcal{CN}(0, \sigma_w^2)$ with uncorrelated real and imaginary components.

The ACF and PSD approximates generated by the above AR model need to be evaluated. This can be done by either comparing against empirical channel measurements taken from the field, or from a specified analytical approach as defined in equations (3.30) and (3.31) above. As the plausibility of the first option is highly dubious due to practical and economical constraints, the second option is employed herein instead. There are a few methods proposed in the literature to adjust the AR model parameters as per the desirable fading covariance statistics. Chief among the various methods is the one that employs Yule-Walker equations [107, 125, 133, 134]. The Yule-Walker approach is considered further here as the viability of other methods is susceptible to extra computational complexities. Let the PSD of AR(L) fading model be defined as below [107, 133, 134]

$$S_{hh}(f) = \frac{\sigma_w^2}{|1 + \sum_{l=1}^{L} a_l \mathrm{e}^{-j2\pi f l}|^2} \tag{3.32}$$

The AR(L) parameters are ruled by the desired ACF model, $R_{hh}(l)$ can be denoted recursively

$$R_{hh}(l) = \begin{cases} -\sum_{m=1}^{L} a_m R_{hh}(l-m) & , \quad l \geq 1 \\ -\sum_{m=1}^{L} a_m R_{hh}(-m) + \sigma_w^2 & , \quad l = 0 \end{cases} \tag{3.33}$$

and in matrix form given by

$$\boldsymbol{R}_{hh}\mathbf{a} = -\boldsymbol{v} \tag{3.34}$$

where

$$\boldsymbol{R}_{hh} = \begin{bmatrix} R_{hh}(0) & R_{hh}(-1) & \dots & R_{hh}(-L+1) \\ R_{hh}(1) & R_{hh}(0) & \dots & R_{hh}(-L+2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{hh}(L-1) & R_{hh}(L-2) & \dots & R_{hh}(0) \end{bmatrix} \tag{3.35a}$$

$$\mathbf{a} = [a_1 \ a_2 \ \dots \ a_L]^{\mathrm{T}} \tag{3.35b}$$

$$\boldsymbol{v} = [R_{hh}(1) \ R_{hh}(2) \ \dots \ R_{hh}(L)]^{\mathrm{T}} \tag{3.35c}$$

$$\sigma_w^2 = R_{hh}(0) + \sum_{l=1}^{L} a_l R_{hh}(-l) \tag{3.35d}$$

Solving the L set of Yule-Walker equations for the desired ACF, the generated AR(L) process yields the following ACF estimate

$$\hat{R}_{hh}(l) = \begin{cases} R_{hh}(l), & 0 \le l \le L \\ -\sum_{m=1}^{L} a_m \hat{R}_{hh}(l-m), & l > L \end{cases} \tag{3.36}$$

To apply the above equations, the $\boldsymbol{R}_{hh}$ is assumed to be of positive definite Toeplitz type.

3.6.3 FSMC Channel Modelling

A conventional 1st-order FSMC model is explored here relying on common aspects in [39, 65, 82, 107, 125-132]. Despite the fact that some of these references have challenged such a scheme, plenty of studies have shown that the 1st-order system is sufficient for adequate and tractable analysis and results. The methodology proposed in [126], which constituted the basis for other studies [39, 65, 82, 107, 127-132], is adopted herein to form a FSMC channel to reflect the Rayleigh fading statistics.

An FSMC channel model is a discrete stochastic process in which the current state depends on the complete history of past states through the most recent state only. Among a few different approaches, it can be built by partitioning γ (the received SNR, whether for bits or symbols) into fixed number of states or intervals. Let us consider an N channel state space $S = \{s_1, \ s_2, \dots, s_N\}$ and their corresponding BER, or crossover, also called transition, probability P_{en} where the subscript 'e' stands for error and

$n \in \{1, 2, \dots, N\}$. If we are using M-ary constellation symbols, which is commonly the case, then these crossover probabilities represent the symbol error rate (SER). Let $P_{n,j}$ be the state transition probability and π_n be the steady state probability such that $\sum_{n=1}^{N} \pi_n = 1$ for the simplest equiprobable SNR quantisation method. The PDF given in (3.27) is revised to consider the AWGN and instantaneous γ, recalling that the fade amplitude is characterised by Rayleigh statistics, while the power of which has an exponential probability distribution given by

$$f(\gamma) = \frac{1}{\bar{\gamma}} e^{-\gamma/\bar{\gamma}} \ , \ \gamma \in [0, \infty) \tag{3.37}$$

where $\bar{\gamma}$ is the average SNR. The 1st-order FSMC constitutes the transitions to happen between adjacent states, and hence the probability $P_{n,j} = \Pr[s_n | s_j] = 0$ if $|n - j| > 1$. Let the boundary separation between different channel states represented by $\Gamma = \{\Gamma_1, \ \Gamma_2, \dots, \Gamma_N\}$, i.e., when the received SNR is in the interval $[\Gamma_n, \Gamma_{n+1})$ at time k, the channel state is defined to be s_n. Then the forward and backward transition probabilities and the level crossing rate (LCR) are given by [39, 65, 125-128, 130], respectively

$$P_{n,n+1} \approx \frac{\Phi(\Gamma_{n+1}) T_s}{\pi_n}, n = 0, 1, \dots, N-1 \tag{3.38}$$

$$P_{n,n-1} \approx \frac{\Phi(\Gamma_n) T_s}{\pi_n} \ , \qquad n = 1, \dots, N \tag{3.39}$$

$$\Phi(\Gamma_n) = \sqrt{\frac{2\pi\Gamma_n}{\bar{\gamma}}} f_D \mathrm{e}^{-\Gamma_n/\bar{\gamma}} \tag{3.40}$$

The above assumes the time axis has "slots" of identical size equal to the symbol period T_s.

3.6.4 Proposed FSMC Approximation Method

There are varieties of statistical approximation methods in the literature. The method that is explored herein has solid foundations in economic, finance, and econometric modelling systems. Tauchen pioneered this field and the work is well known and adopted in many studies. Tauchen contributed to solving functional

equations where the state variables have AR patterns [135, 136]. Computational simplicity and generating almost accurate results under uncorrelated error terms are recognised features attributed to this approximation scheme. It is to the best of this author's knowledge that such an approach has never been addressed or attempted properly in the signal processing field. Hence, this study has contributed as a first lead of deploying Tauchen procedure in the FSMC modelling of wireless fading channels.

To use the approach laid down by Tauchen, one must assume the process values stay within bounded intervals to solve the problem at hand. As stated earlier in this Chapter, these intervals are curbed by N different channel states, which are generated by AR(1) channel model as given in (3.31). These intervals and states are also assumed to be equally spaced to make the Tauchen approach valid. Let the probability of $w(k)$ be such that $P[w(k)] \leq u = F(u/\sigma_w)$, where u is any value and F is the cumulative distribution function (CDF) with unit variance. The following assumptions were made by Tauchen

$$\begin{aligned} s_1 &= -s_N \\ \zeta &= s_n - s_{n-1} \\ s_N &= m\sigma_h = m(\sigma_w^2/(1-a^2))^{1/2} \end{aligned} \tag{3.41}$$

where m is any multiplicity number. There is no particular rule established to set the value for this multiplication parameter, however [135, 136] stated that $1.2 \times ln(N)$ and 3 could be proposed. From this, the transition probabilities for $j \in [2, N-1]$ can be calculated

$$P_{n,j} = F\left(\frac{s_j - as_n + \zeta/2}{\sigma_w}\right) - F\left(\frac{s_j - as_n - \zeta/2}{\sigma_w}\right) \tag{3.42}$$

The above expression can be thought of as the probability that the event $as_n + w \in \left[s_j - \zeta/2\,, s_j + \zeta/2\right]$ takes place. While the transition probability from state n to state 1 is

$$P_{n,1} = F\left(\frac{s_1 - as_n + \zeta/2}{\sigma_w}\right) \tag{3.43}$$

and the transition probability of leaving state n to state N is

$$P_{n,N} = 1 - F\left(\frac{s_N - as_n - \zeta/2}{\sigma_w}\right) \tag{3.44}$$

The above discrete probabilities converge in a weakly sense to their continuous terms in the stochastic recursive model.

3.6.5 Complexity Analysis

The computational complexity of the proposed Tauchen approximation for FSMC channel modelling is assessed with reference to the conventional LCR approach in this section. The computational complexity usually includes the overall operation of mathematical addition, subtraction, multiplication and division procedures. The complexity analysis of FSMC channel models is a threefold challenging task. While assuming the observed samples are statistically independent, the complexity involves the following three steps that need to be efficiently computed: 1) Probability of the observation sequence for a given model, 2) Selection of the corresponding state sequence, and 3) Adjustment of the model parameters. These steps are common in the general context of most HMMs encountered in various applications [137, 138].

As stated earlier, the widely adopted procedure for the iterative estimation of LCR-FSMC parameters is the forward-backward algorithm. Upon giving a finite data sequence of $\forall k \in [1, K]$ samples for training, the forward-backward algorithm smoothly evaluates the likelihood of such data and coordinates sufficient statistics for the FSMC updated parameters according to the Baum-Welch algorithm [137, 138]. Some approximations need to be considered in order to proceed with the computational analysis. The linear expressions given in (3.38) and (3.39) can be applied directly, while the Taylor series expansion is the best candidate to approximate the exponential and square root functions given in (3.40). The CDF of Tauchen-FSMC model expressed in (3.42) can be attended by using the well-known error function $F(z) = erf(z)$. For the purpose of presentation clarity and consistency, and by referring to the relevant functions in [139], these approximations are provided as shown below

$$e^{-z} = \sum_{i=0}^{\infty} \frac{z^i}{i!} = 1 - z + \frac{z^2}{2!} - \cdots \tag{3.45}$$

$$\sqrt{z} = f(z_o) + (z - z_o)\dot{f}(z_o) + (z - z_o)^2\ddot{f}(z_o) + \cdots \tag{3.46}$$

$$erf(z) = \frac{2}{\sqrt{\pi}}\left\{z - \frac{z^3}{3 \times 1!} + \frac{z^5}{5 \times 2!} + \cdots\right\} \tag{3.47}$$

where $f(z_o)$ is the $\sqrt{.}$ at arbitrary constant z_o, $\dot{f}(z_o)$ is the first derivative of $\sqrt{.}$ at z_o and so on to the end of series.

The steady state probabilities π_n, where each denotes that the FSMC attributes originate in state n, are assumed to be fixed, indicating that they remain in their initial conditions. This is valid in the context of WSS processes where the channel parameters are expected to have trivial variations. Furthermore, the common regime of weak channels is addressed by having the value of z to be small enough to make such an assumption admissible. This also is a common practice in most studies since previous results have shown that the most partitioning happens at the regions of low SNR levels where the error probability is significant [127, 131]. Therefore, the first two terms of (3.45) and the second term only of (3.46) are used to approximate the function (3.38)-to-(3.40), while the first term of (3.47) is accounted for by the approximation of (3.42). Successively performing the above procedures over the entire K training population, the time operational requirements $\mathcal{O}(.)$ were developed for both the LCR-FSMC and Tauchen-FSMC modelling schemes, as shown in Table 3.5 below. The parameters that are just calculated once in the Tauchen-FSMC model are discarded from the complexity valuation.

Table 3.5 FSMC complexity operations.

Model	Mul/Div	Add/Sub	Total
LCR-FSMC	$\mathcal{O}(2(N^2 + N)K)$	$\mathcal{O}(4NK)$	$\mathcal{O}(2(N^2 + 3N)K)$
Tauchen-FSMC	$\mathcal{O}(2NK)$	$\mathcal{O}(8NK)$	$\mathcal{O}(10NK)$

Using the partition policy of 10 states, the above total computation figures would be 26×10^6 and 10×10^6 for the LCR and Tauchen FSMC models, respectively, for a 10^5 long training sequence. This is a modest example applicable for Rayleigh faded channels; however, the difference in computational figures becomes exponentially larger as the number of states increases. The situation of extensively large number of

states is expected in diverse applications, such as speech recognition, where the acceptable minimum number of states lies between 32 and 256 for modest performance [137]. Irrespective of the situation, the difference in the time computational loads, and also in memory wise, is undoubtedly apparent in favour of the Tauchen-FSMC model compared to LCR-FSMC.

3.7 Simulation Results and Discussion

Simulation exercises are provided to validate the performance of the proposed Tauchen-FSMC model approximation. An AR(1) model generator based on the JC fading channel is developed first. The simulation scenario assumes a slowly fading channel with one resolvable path and without a LOS component. The channel fading parameters are governed by the values $f_D T_s = 0.01$ and $T_s = 0.1$ ms. This is equivalent to a vehicular mobility of 60 km/hr, which is typically expected as the average speed in urban areas. The SNR is assumed to be 0 dB and the time stream of the generated Rayleigh channel envelope is depicted in Fig. 3.11 below. An AR(1) model is simulated using a = 0.8 as shown in Fig. 3.11, while the associated ACF and the bell shape like PSD trends of which are shown in Figs. 3.12 and 3.13 respectively.

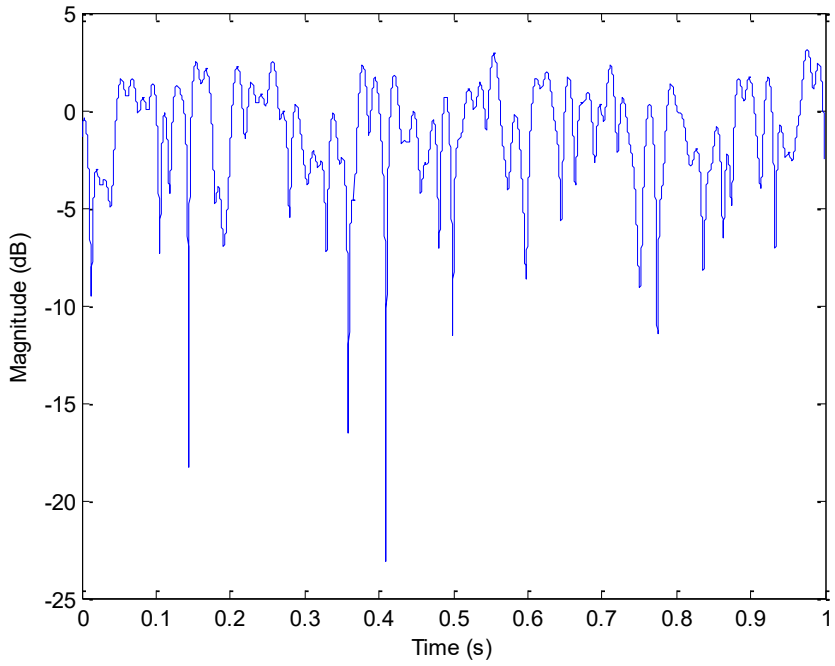


Fig. 3.11 The envelope of AR(1) slow fading Rayleigh channel for a = 0.8.

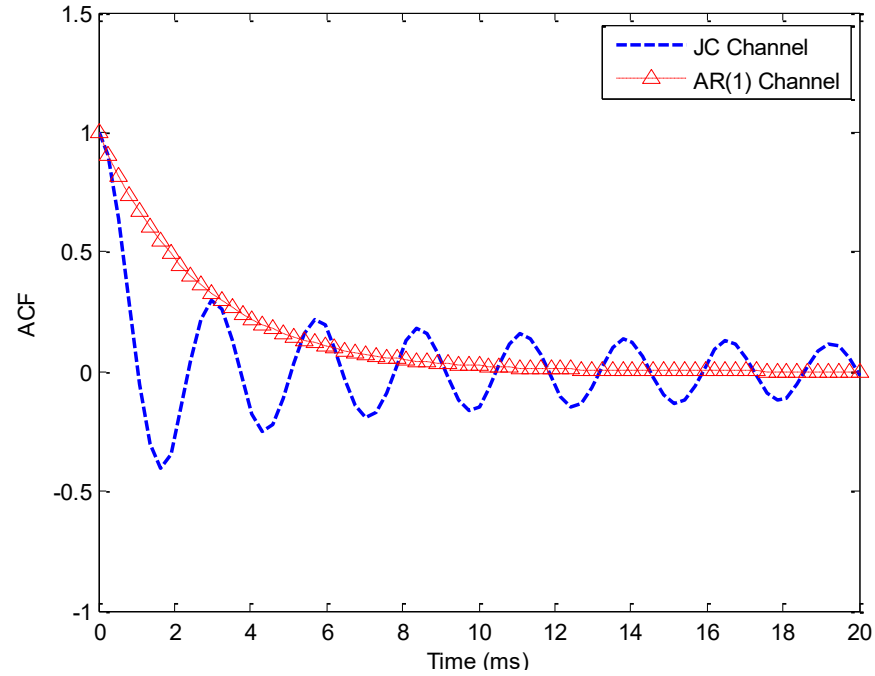


Fig. 3.12 ACF of slow fading Rayleigh channel for $a = 0.8$.

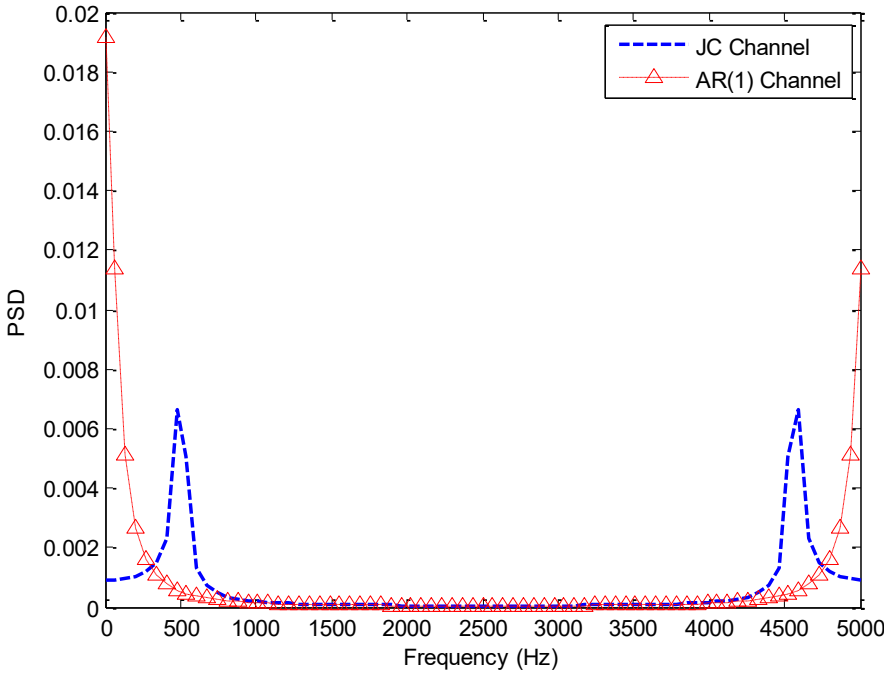


Fig. 3.13 PSD of slow fading Rayleigh channel for $a = 0.8$.

The ACF figure is generated using the Yule-Walker algorithm. Applying the FFT yields the PSD. The Jake and Clark channel model is also depicted in Figs. 3.12 and 3.13 for comparison with the AR(1) results. It is evident that the Jake and Clark model has more resemblance to the Bessel function in the time domain, while it accurately

identifies the Doppler frequency f_D in the frequency domain. On the other hand, the AR(1) is considered as an approximation to such Jake and Clark model behaviour. The AR(1) follows the envelope of Bessel-like function of Jake and Clark model in the time domain, while the PSD of which starts to pick up almost nearby the f_D in the frequency domain. To some extent, such simple approximation could be acceptable in general applications, while when more accurate channel results are required, the order of the autoregressive model needs to be increased. This agrees with suggestions in some studies to increase the order and make it tens, or a few hundred, to fit particular applications. However, this will be at the cost of extra computational loads.

Next, the simulation of 1st-order FSMC channel models is considered. The channel fading envelope, or its associated SNR, is portioned into 10 equal intervals. Firstly, the LCR method and the transition probabilities are analysed. Their cumulative trend and LCR curve are calculated, as given earlier, the results of which are depicted in Figs. 3.14, 3.15, and 3.16, respectively. As it was expected, the transition probabilities and their cumulative trends show sharp crossovers between adjacent states and with very small probabilities to transit to other far states. The footprint contours exhibit heavy concentrations of states transiting to themselves or to their neighbouring states. This confirms the applicability of this statistical model as claimed by most prior studies.

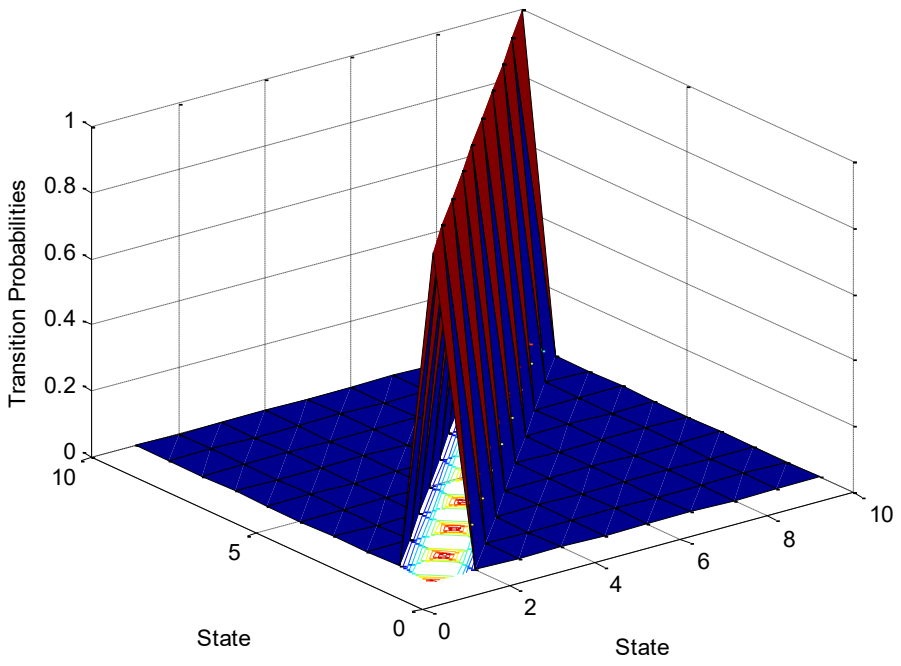


Fig. 3.14 Transition probabilities of LCR-FSMC channel.

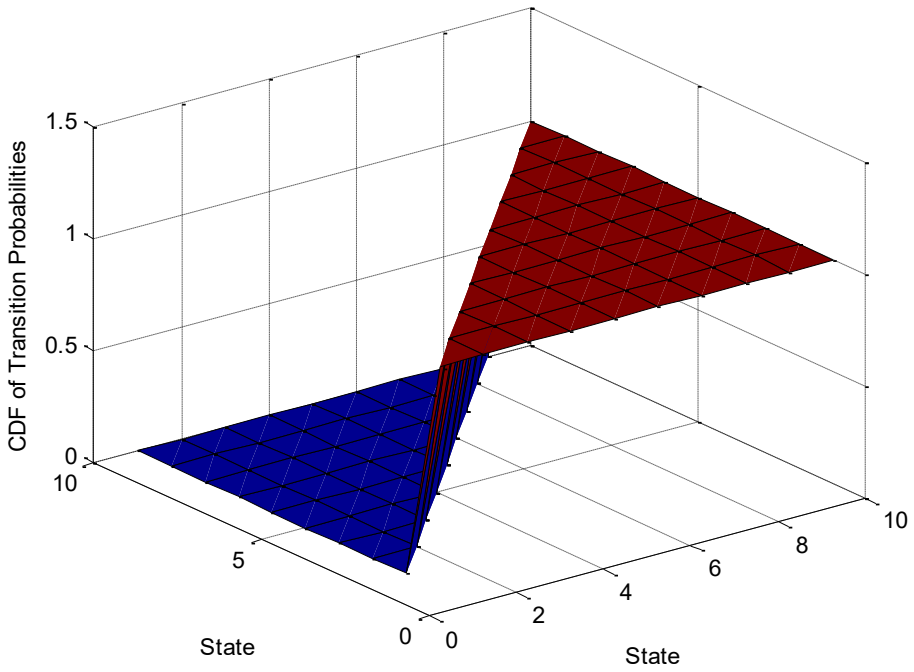


Fig. 3.15 Cumulative transition probabilities of LCR-FSMC channel.

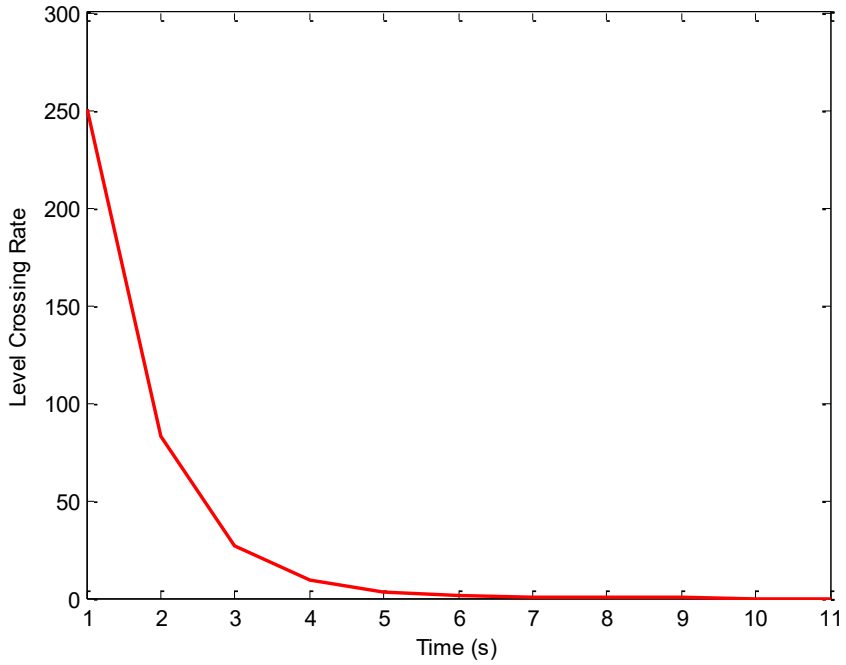


Fig. 3.16 LCR trend of LCR-FSMC channel.

Fig. 3.16 shows a persistence rate for a state to remain or to cross to the next neighbouring states only. This rapidly declining exponential curve obviously depicts the state's mood for making short-time traversing rather than long time traveling to distant states. If more state intervals are considered, the sharper is this exponential behaviour,

and hence consolidate shorter transition paths. This suggests to have a larger pool of state partitions. However, there is no feasible tool to examine the influence of the envelope fading parameter a on the LCR-FSMC model. Hence, this model can be identified as being insensitive to fading strength variations, which is one of the main findings of this Chapter.

Secondly, the Tauchen approximation method is examined for the 1st-order FSMC channel model. Figs. 3.17 and 3.18 below illustrate the results implemented for the same envelope parameter a = 0.8 as before. Despite the results showing almost comparable statistical channel features as that of the LCR method, a more resilient behaviour can now be detected. This is mainly attributed to the direct influence of the a parameter on the AR filter bandwidth. In other words, such an effect is explicitly interpreted in terms of the channel memory and its profound reliance on old or current states. The smaller the bandwidth, the stronger the link is to current states. The converse is also true and hence there would be more correlation governing remote jumps into past or future channel states.

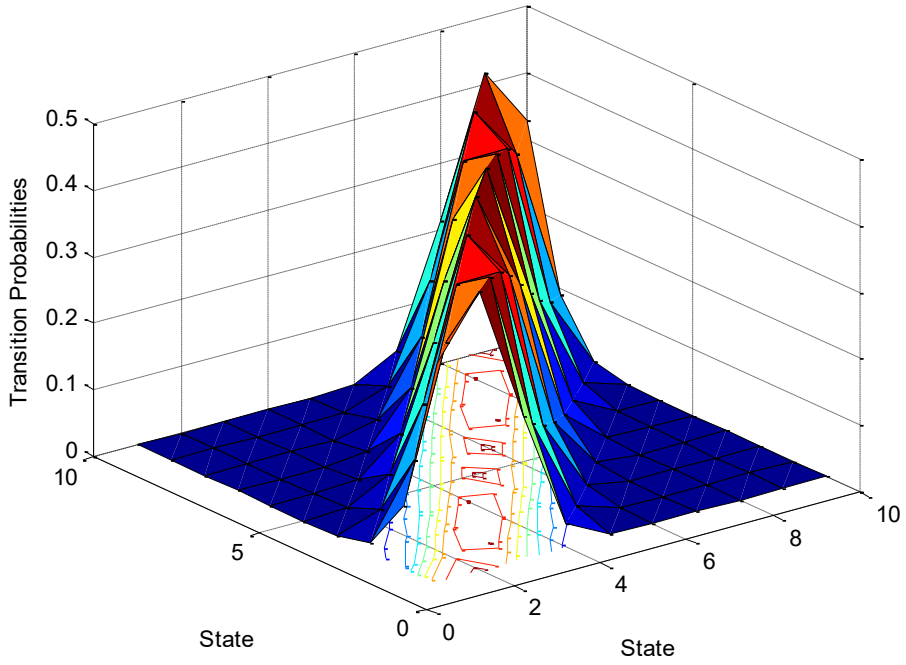


Fig. 3.17 Transition probabilities of Tauchen-FSMC channel for a = 0.8.

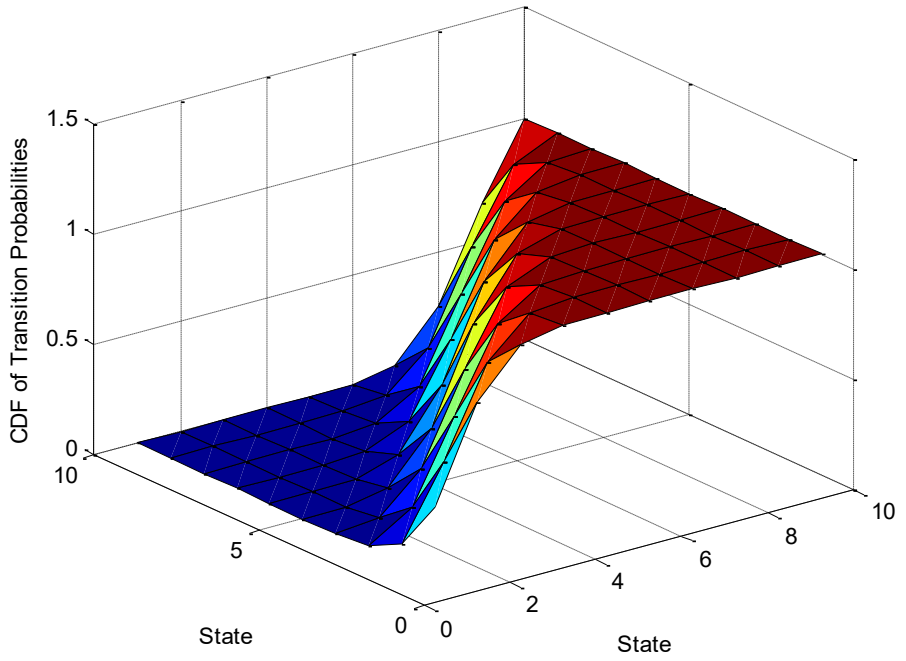


Fig. 3.18 Cumulative transition probabilities of Tauchen-FSMC channel for $a = 0.8$.

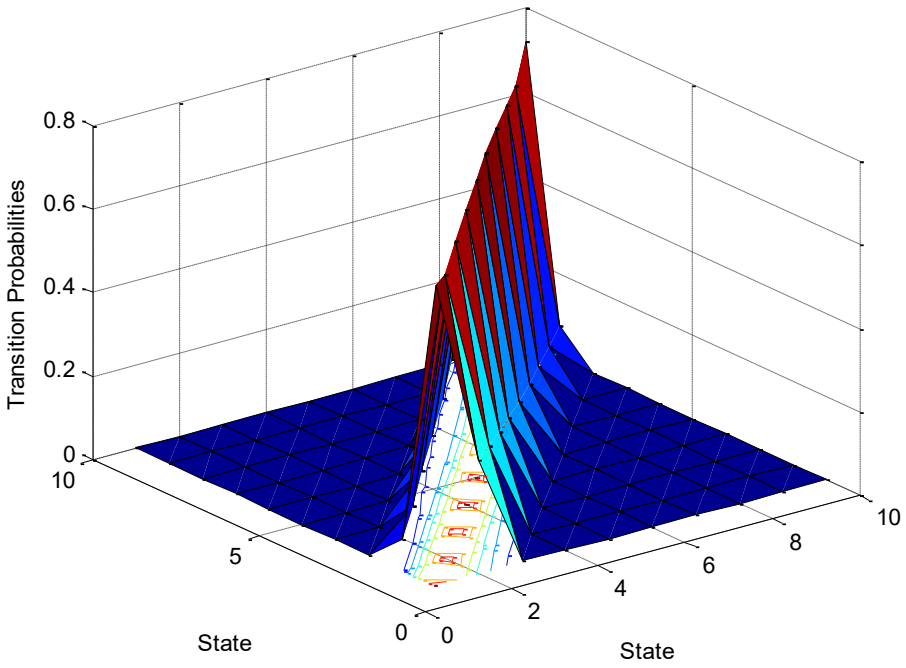


Fig. 3.19 Transition probabilities of Tauchen-FSMC channel with $a = 0.95$.

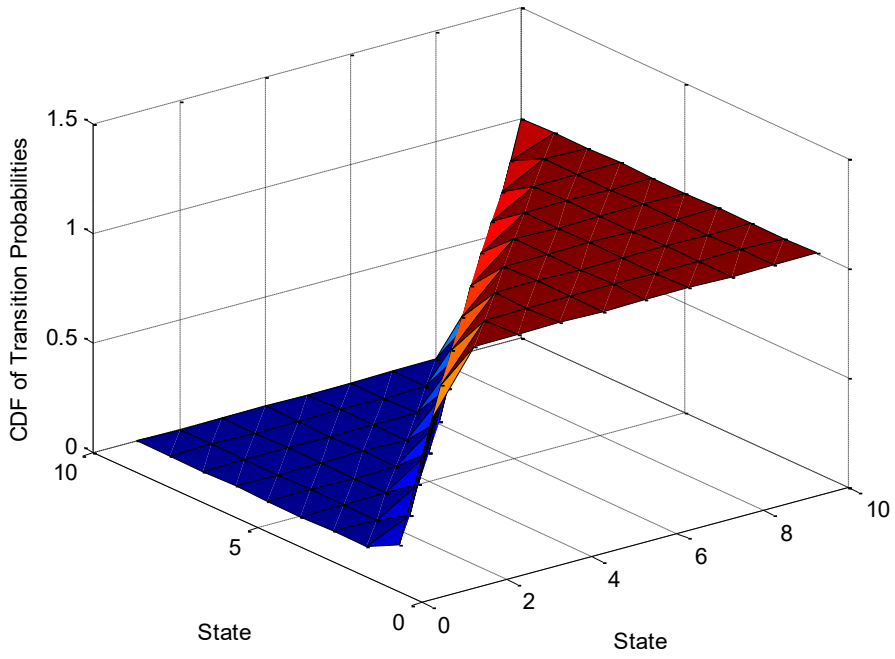


Fig. 3.20 Cumulative transition probabilities of Tauchen-FSMC channel with a = 0.95.

Another envelope parameter value of a = 0.95 is examined to consolidate the above important finding and the generated results are illustrated in Figs. 3.19 and 3.20. The transition probabilities and their cumulative trends show more resemblance to those obtained for the LCR method. Such statistical behaviour reflects on the flexibility of the Tauchen model in determining the AR system order on the expressed statistical approximations. This is quite different to what was experienced [126, 127, 130] with respect to the static behaviour of the LCR method, which hence can be considered as a merited feature counted for the Tauchen method.

3.8 Summary

This chapter has provided some insights on the fundamental tools and techniques necessary for the analysis and developments in the remaining chapters of this work. These include M-ary digital signalling, multiple antennas and carriers, combing and spatial-time configurations, along with some guidelines on their performance metrics. Spatiotemporal coding methods have also been explained, and generally the schemes that enjoy space and time diversity have revealed unrivalled levels of improvement compared to their counterparts without such features. Details on the modelling of wireless fading channels, particularly of the Rayleigh type, using 1st-order AR and

FSMC schemes have also been elaborated. A novel approach for approximating these models using Tauchen method, which is borrowed from econometric studies and applied for fading channel approximation for the first time in this work, has also been presented. The performance results have been proved very competitive and hold more promises than conventional methods in this regard. This accordingly represents one of the prime achievements of the work of this study and upon which further developments can be attempted in future to expand Tauchen method and include other types of fading channels.

Chapter 4

ADAPTIVE CHANNEL ESTIMATION

4.1 Introduction

Guided by the facts delivered in the previous chapters, the indispensable role of channel knowledge to the overall performance of wireless communication systems, and particularly CR systems, has become well discerned. It has been shown that the asset of meaningful transmission channel constitutes a crucial part of any successful SS strategy in CR applications [1]. The SS reliability is challenged by channel uncertainty, since CRs have to increase their sensitivity in order to be able to distinguish between faded or shadowed PUs' signals from white spaces [89]. Such channel, or CSI, knowledge is also considered significant to ACM in a wide range of wireless communication systems and CR systems [43]. Obviously, the absence of such knowledge rendered the blind signals identification a very difficult task to achieve [59, 60]. With time, the emphasis of CSI estimate and feedback provisions is unfolded and becomes very explicit [140, 141]. Keeping an eye on the precious transmission bandwidths, the feedback link must be of low-rate, quantised and honoured in the SUs' bands [1, 140, 141].

Inspired by the above, this chapter is intended to furtherly elaborate the analyses of mostly recognised CSI estimators accessible in the literature. As such, the LS, MMSE and ML estimation methods are given direct attention. The optimum state-space estimation using KF is also examined herein. The consolidated IMM and KF algorithms are for the first time explored for the purpose of CSI estimation in this work. The IMM algorithm is well known in tracking manoeuvring targets and it is interfaced with KF for the purpose of CSI estimation herein. As previous encounters of which cannot be clearly identified in the literature, this is advocated as an empowered adaptation enabling and thereby manifests another valuable finding among others of this study. Another milestone achievement added in this chapter is by proposing a new paradigm of applying the IMM-KF adaptation policy directly on the eigen channels disintegrated using the SVD technique.

4.2 Rudimentary Estimation Theory

The principles of estimation theory are not new; they mainly started to receive high interest in the late 1930s and early 1940s when the breakthrough of time series analysis explicitly addressed by H. Wold and A. Kolmogorov, and particularly established by the most famous work of N. Wiener [142, 143]. Estimation theory is an engineering term, which is usually named as parameter estimation or point estimation by statisticians. It can be found in diverse fields, such as communications, control, econometrics, and signal processing and all share the common demand of wanting to estimate the values of a collection of desired parameters. In other words, the problem is one of obtaining measurements and from which estimate the numerical value of a real or complex vector that describes the system under study. If the vector is discrete in a finite-dimensional space, then there is no difference between estimation theory and detection theory [142, 143].

Observations are typically arranged in strings or arrays. The commonality is to use vectors in reference of the strings, and matrices in reference of the arrays. In order to determine good estimators, the first step is to put the observed data in a suitable mathematical model. Since the data are inherently of random nature, a PDF is usually employed to describe them. That is for the signal model given in (3.1) and (3.2), the PDF $p([y(1)\ y(2)\ \ldots\ y(K)]^{\mathrm{T}};\varphi)$ is parametrised by the unknown parameter φ, i.e., there is a class of PDFs where each one is different due to a different value of φ. Intuitively, because φ affects the probabilities of φ, the value of φ should be inferred from the observation set of $\boldsymbol{y}$. Suppose any combination of the unknowns is given in a vector form $\boldsymbol{\varphi} = [h\ x\ \sigma_n^2]^{\mathrm{T}}$, then the multi-variate normal PDF is denoted by [142, 143]

$$p(\boldsymbol{y};\boldsymbol{\varphi}) = \frac{1}{(2\pi\sigma_n^2)^{\frac{K}{2}}} exp\left[-\frac{1}{2\sigma_n^2}\sum_{k=1}^{K}(y(k) - hx(k))^2\right] \tag{4.1}$$

From the above, it is obvious that the performance of any estimator is critically dependent on the PDF assumptions. An estimator is called robust if slight changes in the PDF do not lead to severe deterioration in the estimator performance. Furthermore, estimators of the type given in (4.1) belong to the family generally known as classical, and the unknown parameters of interest are assumed to be deterministic.

In the Bayesian estimation, *a priori* knowledge on the unknowns is incorporated. The parameters that are attempted for estimation are hence viewed as a realisation of the RV $\boldsymbol{\varphi}$. As such, the joint PDF is used to describe the data [142, 143]

$$p(\boldsymbol{y};\boldsymbol{\varphi}) = p(\boldsymbol{y}|\boldsymbol{\varphi})p(\boldsymbol{\varphi}) \tag{4.2}$$

where $p(\boldsymbol{\varphi})$ is the prior PDF, indicating the knowledge about $\boldsymbol{\varphi}$ before any data are observed, and $p(\boldsymbol{y}|\boldsymbol{\varphi})$ is a conditional PDF, indicating the knowledge provided by the data $\boldsymbol{y}$ conditioned on knowing $\boldsymbol{\varphi}$. Once the PDF has been defined, the problem reduces to one of determining an optimal estimator or function of the data. An estimator hence may be thought of as a regime that assigns values to $\boldsymbol{\varphi}$ for each realisation of $\boldsymbol{y}$. Some concepts of the linear estimation essential to this work are presented in what follows.

4.2.1 Unbiased Estimator

For an estimator to be unbiased, it means on average the estimator yields the true value of the unknown parameter. Since the range of the unknown parameter may be generally within the range $a < \varphi < b$, unbiasedness asserts that irrespective of the true value of φ, the selected estimator yields the value of unknown parameter on an average basis. Mathematically, that means the following condition is to be met [142, 143]

$$\mathbb{E}(\hat{\varphi}) = \varphi \quad , \quad \forall \varphi \in (a, b) \tag{4.3}$$

where (a, b) denotes the possible range of φ and the hat in $\hat{\varphi}$ is the estimated value. Although not necessary, but most often, unbiased estimators tend to have symmetric PDFs centred about the true value of φ. Generally, an unbiased estimator does not necessarily mean it is a good estimator. Conversely, biased estimators are poorly characterised by a persistent systematic error, which presumably should not be present.

4.2.2 Minimum Variance Unbiased Estimator

In a search for optimal estimators there is a need to adopt some optimality criterion. A natural choice is the MMSE, which is defined to minimise the following [142, 143]

$$\mathrm{MSE}(\hat{\varphi}) = \mathbb{E}[(\hat{\varphi} - \varphi)^2] = var(\hat{\varphi}) + bias^2(\varphi) \tag{4.4}$$

which shows that the MSE is composed of two error terms, the first is owed to the variance of the estimator and the second is related to the bias of the estimator.

Practically, the MMSE seems to be unfeasible since any criterion that depends on the bias usually leads to an unrealizable estimator. An alternative approach is to restrict the bias to be zero and search for an estimator that minimises the variance only. Such an estimator is called the minimum variance unbiased estimator (MVUE), which is more common in reality. A MVUE has the effect of focusing the PDF of the estimation error about zero and the probability of which is unlikely to be large [142, 143].

Although a MVUE exists, it may be challenging to find one. That is because there is no straightforward procedure to always rely on while designing such an estimator [143]. Along such a quest, some useful steps can be stressed as below [142, 143]:

- Check if an estimator satisfies the Cramer-Rao lower bound (CRLB) constraint;
- Apply the Rao-Blackwell-Lehmann-Scheffe (RBLS) theorem;
- Not only restrict the class of estimators to be unbiased, but also make them linear.

4.2.3 Cramer-Rao Lower Bound

Assume that the PDF $p(y;\varphi)$ fulfils the following condition for all φ [142, 143]

$$\mathbb{E}\left[\frac{\partial ln\big(p(\boldsymbol{y};\varphi)\big)}{\partial\varphi}\right]=0 \tag{4.5}$$

where the $\mathbb{E}[.]$ is over $p(\boldsymbol{y};\varphi)$. The variance of an unbiased estimator must satisfy

$$var(\hat{\varphi})\geq\frac{1}{-\mathbb{E}\left[\frac{\partial^2 ln\big(p(\boldsymbol{y};\varphi)\big)}{\partial\varphi^2}\right]} \tag{4.6}$$

where the derivative is evaluated at the true value of φ and the expectation is over $p(\boldsymbol{y};\varphi)$. Then, an estimator may be found to attain the bound for all φ if and only if

$$\left[\frac{\partial ln\big(p(\boldsymbol{y};\varphi)\big)}{\partial\varphi}\right]=f(\varphi)(g(\boldsymbol{y})-\varphi) \tag{4.7}$$

for some functions $g(.)$ and $f(.)$. That estimator is a MVUE with $\hat{\varphi} = g(\boldsymbol{y})$ and the minimum variance of which is $1/f(\varphi)$.

The CRLB allows to determine that for any unbiased estimator the variance must be greater than or equal to a given value. If an estimator exists whose variance equals the CRLB for each value of φ, then it must be the MVUE and, as such, the theory of the CRLB straightforwardly yields the estimator. If an estimator whose variance equals the bound may happen not to exist, yet a MVUE may still exist. The RBLS theorem needs to be sought per se, to first find a sufficient statistic, one which uses all the data sufficiently, and then find a function of the sufficient statistic, which retains an unbiased estimator of φ.

4.2.4 Rao-Blackwell-Lehmann-Scheffe Theorem

The term of sufficient statistics, which play a crucial role in both the estimation and detection theories, needs to be defined first. Using Neyman-Fisher factorisation, the PDF $p(y;\varphi)$ can be revised as follows [142, 143]

$$p(\boldsymbol{y};\varphi) = g(T(\boldsymbol{y}),\varphi)f(\varphi) \tag{4.8}$$

where $g(.)$ is a function depending on $\boldsymbol{y}$ only through $T(\boldsymbol{y})$ and $f(.)$ Is a function depending on $\boldsymbol{y}$, then $T(\boldsymbol{y})$ is a sufficient statistic for φ. Alternatively, if $T(\boldsymbol{y})$ is a sufficient statistic for φ, then the PDF can be factorised as given in (4.8).

So much as that, the RBLS can now be stated as follows [142, 143]. If $\breve{\varphi}$ is an unbiased estimator of φ and $T(\boldsymbol{y})$ is a sufficient statistic for φ, then $\hat{\varphi} = \mathbb{E}\big(\breve{\varphi}|T(y)\big)$ is a valid estimator of φ, not dependent on φ, unbiased and of lesser or equal variance than that of $\breve{\varphi}$ for all φ. If the sufficient statistic is complete, then $\breve{\varphi}$ is a MVUE. This requires the estimator to be linear, or more restrictively to be best linear unbiased estimator (BLUE). When the data are truly Gaussian, then the BLUE is MVUE.

4.3 Channel Estimation Techniques

The estimation of CSI parameters constitutes an important task in most classical and modern wireless communication systems, and predominantly in CR environments. That

is, a resilient performance planned for the best practices of spectrum sustainable access and utilisation cannot be attained without knowing the CSI perfectly. In the absence of having the full knowledge of CSI, some operational efficacies in wireless communication systems and CR systems might be challenged. That is because most of advanced wireless technologies essentially rely on the instantaneous CSI being accurately acquired. Given the time-varying nature of wireless channels, the assumption of having the full knowledge of CSI accessible at both the transmitter and the receiver is also not practical. Therefore, the design and development of robust and efficient CSI estimation techniques is inevitable for wireless systems to achieve the desirable QoS and operational levels [44-46, 60, 112, 113].

A great deal of research studies on CSI estimation techniques can be identified over the past decades, and only few notable are itemised here [46, 112, 113, 144-147]. These techniques have a long and rich history in single transmission entities [46, 144], and the same techniques can be directly applied for multiple carrier and multiple antenna communication systems [112, 113, 145, 146]. Generally, CSI estimation strategies can be broadly categorised into two main groups: the first is called non-blind, supervised or trained; while the second group is called blind, unsupervised or untrained. The non-blind CSI estimation is the most common method in wireless communication systems due to its flexibility and less complexity. It also can be divided into two main subgroups, namely; data aided (DA) and decision directed channel estimation (DDCE). The DA can be easily realised by having particular training sequences periodically inserted into the data frames in the time or frequency domain. By being supplied in advance with *a priori* knowledge of the training sequences, the receiver can hence readily estimate the CSI parameters. This method is typical to OFDM systems and known as pilot-assisted (or aided) channel estimation (PACE) [112, 145] or pilot symbol aided modulation (PSAM) [48, 147]. Whereas the DDCE method arranges all training symbols at the first OFDM blocks. The CSI consistent with the training blocks are first estimated, and that consistent with the subsequent data blocks can be tracked with further gain supplied by the demodulated data [46, 112, 113].

However, the employment of the training sequences degrades the system performance in terms of SE and PE as additional spectrum resources are demanded for the training transmission. Therefore, the blind CSI estimation techniques were introduced as an alternative option for the CSI estimation by using the intrinsic features

of data signal itself instead of pilot signal [146]. Despite their high SE, these techniques might suffer from other issues related to slow convergence speed, high computation complexity, and sensitivity to noise, which may impose some limits on their practical implementations. The work of this study adopts the approach of non-data aided (NDA) CSI. Among several estimators, those described in what follows are intrinsic to a wide range of wireless systems, yet, they can be equally likely applied in blind or non-blind situations.

4.3.1 Least Square

The LS estimator (LSE) is the simplest form among estimators. It is usually applied in situations where precise knowledge of the data statistical characteristics is not provided. It is also recommended when an optimal estimator cannot be found or may be too intricate to realise in practice [142, 143]. Without statistical knowledge, the CSI can be treated as deterministic but unknown. In this case, the LSE can be readily exploited as it requires no statistical information about the channel or data [143, 144, 145]. It is also equally valid for AWGN or non-AWGN, or when a training sequence is available or not. For the linear model given in (3.2), the LSE minimises [142, 143, 144, 145]

$$\begin{aligned} \hat{h}_{LS} &= argmin_h(\boldsymbol{y} - h\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y} - h\boldsymbol{x}) \\ &= argmin_h\|\boldsymbol{y} - h\boldsymbol{x}\|^2 \end{aligned} \tag{4.9}$$

By differentiating with respect to each channel coefficient and setting the result to zero, the following closed form expression for the LSE can be obtained [46, 113, 142, 143, 144]

$$\hat{h}_{LS} = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} \tag{4.10}$$

The matrix result of $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ multiplication is called Grammian or Gram matrix, and if it is singular then the solution for $\hat{h}_{LS}$ is unique, otherwise many solutions may exist.

4.3.2 Maximum Likelihood

Despite the principle of ML being deceptively simple, it could be inaccessible for some uncertain users. Such principle can be directly applied to a wide variety of

inference problems, and, largely, the results are worthwhile [142]. In many cases, however, the MLEs can be reduced to their equivalent LSEs and MVUEs [142, 143]. Similar to LSE, the MLE also assumes that the CSI is a deterministic quantity with unknown parameters, and it is expected to perform optimally and achieves the CRLB. With MLE, the likelihood of the received data samples given the channel coefficients is maximised. The MLE problem can be stated as follows [142-144]

$$\hat{h}_{ML} = argmax_h p(\boldsymbol{y}|h) \tag{4.11}$$

and for complex AWGN, the ML solution can be obtained by maximising the log-likelihood function (LLF), which relates to the noise statistics, as follows

$$\hat{h}_{ML} = argmax_h[(\boldsymbol{y} - h\boldsymbol{x})^{\mathrm{T}}\boldsymbol{R}_n^{-1}(\boldsymbol{y} - h\boldsymbol{x})] \tag{4.12}$$

where $\boldsymbol{R}_n$ is the $K \times K$ noise covariance matrix over K-length observation samples. Setting the derivative with respect to h equal to zero yields the following simpler form of solution [46, 142-144]

$$\hat{h}_{ML} = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{R}_n^{-1}\boldsymbol{x})^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{R}_n^{-1}\boldsymbol{y} \tag{4.13}$$

The MSE matrix is given by $(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{R}_n^{-1}\boldsymbol{x})^{-1}$. Assuming uncorrelated white noise, it can be easily deduced that the term $\boldsymbol{R}_n$ reduces to $\sigma_n^2\boldsymbol{I}$, relevant factors cancel each other and (4.13) simplifies to the LSE solution given in (4.10) [46, 112, 142-144].

4.3.3 Minimum Mean Square Error

The MMSE estimator is fairly adopted in broad wireless communication systems, especially those employing multi-carrier signalling like the OFDM scheme [46, 112, 144, 145, 147]. Stemmed from its direct adherence to the maximisation of information-theoretic channel capacity, the MMSE estimator can also be found in the core of ACM systems tasked for spectral efficiency perfection [44, 45].

Unlike LSE and MLE, the MMSE is a member of Bayesian estimators' class that exploits *a priori* knowledge to produce more accurate estimation, which is a fundamental rule of estimation theory [142, 143]. They are commonly called as MAP estimators, or MAPEs. As such, the channel coefficients are treated as random variables

with known statistics and the likelihood of which is maximised. The MMSE estimator is designed to minimise the Bayesian MSE cost function as below [142-145, 147]

$$\begin{aligned}\hat{h}_{MMSE} &= argmin_h[\mathbb{E}\{(\boldsymbol{y}-h\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y}-h\boldsymbol{x})\}] \\ &= argmin_h[\mathbb{E}\{\|\boldsymbol{y}-h\boldsymbol{x}\|^2\}]\end{aligned} \tag{4.14}$$

where the expectation $\mathbb{E}(.)$ is evaluated using the joint PDF described earlier in (4.1) and (4.2). Assuming h is of zero-mean and covariance $\boldsymbol{R}_h$, this can be simplified to

$$\hat{h}_{MMSE} = \boldsymbol{R}_h\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{x}\boldsymbol{R}_h\boldsymbol{x}^{\mathrm{T}} + \boldsymbol{R}_n)^{-1}\boldsymbol{y} \tag{4.15}$$

This is exactly the same expression for MAPE providing the assumption that the channel coefficients are Gaussian. Putting this in terms of the LSE yields [46, 112, 142, 143, 145]

$$\hat{h}_{MMSE} = \boldsymbol{R}_h(\boldsymbol{R}_h + \boldsymbol{I}/\gamma_s)^{-1}\hat{h}_{LS} \tag{4.16}$$

It is worth to note an important matter related to the performance of the MMSE estimator. The MMSE estimator performs optimal given the constraint of a linear statistical expectation based on the marginal PDF of h and y. This in turn labels such optimal linear estimator as LMMSE [142, 143] and to find one is not so hard in practice. However, a suboptimal LMMSE estimator is also quite useful, widely in existence and shows acceptable performance in reality. It is usually given in a closed form governed by the first two moments only (means and covariance).

Although the LMMSE estimator improves the performance significantly, the complexity is also increased due to the need to perform a matrix inversion. Therefore, to decide which estimator to proceed with is always a trade-off between minimising overhead and optimizing channel estimation performance [142-144].

4.3.4 State-Space Estimation Using Kalman Filter

Owing to its rich versatility, the celebrated KF is pivotal in a wide choice of signal processing and dynamical systems [142, 143]. In the framework of wireless communication systems, KF lends itself very readily for direct channel applications such as equalisation [39, 107], estimation [40] and tracking [118, 144]. The main

requirement for the performance of KFs to be optimal is by having them completely devised under linear Gaussian perturbations. If such a condition is not met, then other versions of sub-optimal KFs can be sought, such as by applying the linearization process to produce an extended KF (EKF) [142, 143] or comparably rerouting to a particle filter [1, 148] or a mixture of both [148].

The KF algorithm may be thought of as a sequential MMSE estimator of a signal embedded in noise, where the signal characteristics are evolving as per a dynamical or state model [142, 143]. It also can be viewed as a recursive LS (RLS) estimator subject to certain initialisation attributes having been taken a good care of [39, 107], while certain approximations render KF equivalent to the least mean square (LMS) algorithm that can be evoked recursively too [144]. Without loss of generality, KFs can come in two different forms, either scalar or vectorised [142, 143], any one of which can be exposed to MIMO, OFDM or combination of both signalling schemes [149-151]. They can be supportive by combining with neural networks to train learning sequences [152] and with beamforming of massive MIMO schemes [153]. They also have been proved to be very useful in other types of wireless settings such as wireless sensor networks [154], just to enumerate a few among many others.

In all of the above studies, an appropriate state-space model for the problem at hand needs to be considered. Such a task can be usually attained by invoking the effective Gauss-Markov model to define the dynamical system features that evolve with time [142, 143]. In the context of fading channels, fortunately, the AR model in (3.31) lends itself to a state-space representation, which enables the application of KF for tracking the time-varying channel. Revisit the expression to have a AR(1) model given below

$$h(k) = h(k-l) + w(k) \tag{4.17}$$

To a greater extent, this AR(1) model is accepted as a good educated guess of the casual channel variations branded by Rayleigh statistics [1, 125, 133, 134, 142-144, 148-154].

Another model is necessary to proceed with the building of KF algorithm. This model is related to the observation of received data sequences, which was explicitly expressed in (3.1) or (3.2) and repeated here for notional convenience

$$y(k) = h(k)x(k) + n(k) \tag{4.18}$$

All variables are naturally assumed to be *iid* and the noise terms are given by $\mathcal{CN}\sim(0,\sigma_w^2)$ and $\mathcal{CN}\sim(0,\sigma_n^2)$ for the state model and observation model, $w(.)$ and $n(.)$, respectively.

The channel state-space and observation models given in (4.17) and (4.18) are now supplying enough details to develop the recursive KF algorithms. Given a discrete time variable is denoted as subscript for ease, the following definitions are due; $\hat{h}_{k|k-1}$ and $\hat{h}_{k|k}$ for the state prediction and estimate at step k, and $P_{k|k-1} \triangleq \mathbb{E}\left\{\left(h_{k|k}-\hat{h}_{k|k-1}\right)^2\right\}$ and $P_{k|k} \triangleq \mathbb{E}\left\{\left(h_{k|k}-\hat{h}_{k|k}\right)^2\right\}$ for the covariance of the prediction and estimation errors, respectively. Assuming the incoming signal sequence has $\mathbb{E}\{|x(.)|^2\}=\sigma_x^2=1$, the recursive algorithms for scalar KF can then be devised as follows [142-144, 148-154]

- Predictions

$$\begin{cases}\hat{h}_{k|k-1}=\hat{h}_{k-1|k-1}\\ P_{k|k-1}=P_{k-1|k-1}+\sigma_w^2\end{cases} \tag{4.19}$$

- Updates

$$\begin{cases}\hat{h}_{k|k}=\hat{h}_{k|k-1}+G_k z_k\\ G_k \triangleq P_{k|k-1}{S_k}^{-1}\\ P_{k|k}=(1-G_k)P_{k|k-1}\end{cases} \tag{4.20}$$

where G_k is the KF gain at step k, the residual sequence $z_k=\left(y_k-\hat{h}_{k|k-1}\right)$ is the residual (innovation) process of zero-mean and covariance $S_k=P_{k-1|k-1}+\sigma_w^2$. For the purpose of initialisation, the values $\hat{h}_{0|0}>0$ and $P_{1|0}\geq 1$ can be arbitrarily selected. A schematic diagram of the recursive KF is portrayed in Fig. 4.1 below [40, 142, 143, 154]. The above KF algorithms can be easily extended to consider the MIMO and OFDM systems by simply converting all scalar variables into their vector or matrix counterparts.

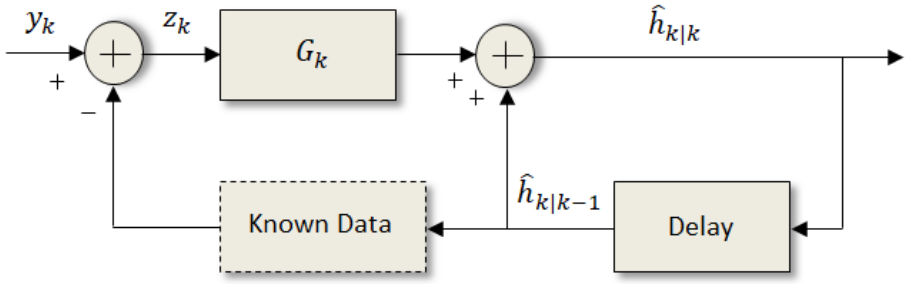


Fig. 4.1 Simplified KF for channel estimation.

The iteration of KF filter continues sample by sample over all received data sequences until convergence is achieved, i.e., until the MSE represented by the expected value of residual covariance $\mathbb{E}(S_k)$ is at its minimum value. It is also noteworthy to bring about the assumption stated above in relation to the variance of the incoming data sequence, which made the above KF algorithms and schematic figure valid. Such a scenario is similar to that of having a known training or data sequence and being supplied to the estimator, as indicated by the dotted block in Fig. 4.1 [40]. Otherwise, the incoming data sequence needs to be suitably detected in conjunction with the channel estimation as will be explored further in the subsequent chapter on AMC systems.

4.4 Interacting Multiple Model-Kalman Filter

The IMM scheme is a member of a big class called Jump Markov Linear Systems (JMLSs), or also dubbed Markov Jump Linear Systems [155, 156]. The versatile applications of JMLSs have been widely endeavoured in a large variety of fields, such as; aerospace, navigation, signal processing, cybernetics, fault diagnoses, control and communication systems. As a reminiscent of such JMLS hybrid systems, IMM forms a linear system whose parameters evolve with time according to the realisation of an FSMC process. The IMM approach, as one of the best suboptimal state-space estimation methods, is widely used for JMLSs. Built on sophisticated hypothesis pruning and merging techniques; IMM offers exceptional estimate performance with low-cost computations. The generic concept of IMM was first heavily investigated in the area of adaptive tracking of manoeuvring targets, and as such insightful surveys can be consulted accordingly [157].

The IMM estimator has an attractive feature attributed to the assumption of having a bank of multiple models working in tandem to track different aspects of the channel state evolution. These models are separate and have dissimilar settings to cope with a variety of possible channel behaviours. The IMM then uses a soft decision by applying a probabilistic combination of the individual estimates at the output of each filter. This is done without making any hard decision as to which model is in effect at a particular time. Another important feature of the IMM estimator lies in the dynamical weighing of the combining probabilities that evolves based on the likelihood function resulting from each individual model. Putting this in the context of IMM-KF amalgamation for CSI estimation, the input to a single KF at time $k+1$ is the estimate $\hat{h}_{k|k}$ at time k. While in the overall IMM-KF build, the input to each KF at time $k+1$ is a probabilistic combination of the k^{th} estimate of all other KFs, thus accounting for all possible model transitions developed for from time to time [155-157]. The IMM-KF has not been expressively attempted for CSI estimation in the literature before, and thus proposed here to drive for further adaptation enrichment.

So much as that given, scalar KFs can be arbitrated for a span of discrete channel states and hence facilitate the IMM algorithm appropriately. For a $\mathcal{J}$-bank of KFs, this algorithm comprises the following core processing steps [155-157].

- Interaction

  Individual KFs are mixed according to the predicted model probabilities at time k. The predicted model probability is given in terms of the model probability in the previous time cycle $\mu^{i}_{k-1|k-1}$ and the channel transition probability occurs from state i to state j, which is denoted by $\pi^{i,j}_{k}$

$$\mu^{i|j}_{k-1|k-1} = \frac{\pi^{i,j}_{k}\mu^{i}_{k-1|k-1}}{\mu^{j}_{k|k-1}} \tag{4.21}$$

  being the conditional model probability, given the channel is in state i and that the transition occurred from state j involving all individual KFs

$$\mu^{i}_{k-1|k-1} = \sum_{j=1}^{\mathcal{J}} \pi^{i,j}_{k}\mu^{j}_{k-1|k-1} \tag{4.22}$$

and the mixing of the state estimates $\bar{h}^i_{k-1|k-1}$ and covariances $\bar{P}^i_{k-1|k-1}$ are given by

$$\bar{h}^i_{k-1|k-1} = \sum_{j=1}^{\mathcal{J}} \mu^{i|j}_{k-1|k-1} \hat{h}^i_{k-1|k-1} \tag{4.23}$$

$$\bar{P}^i_{k-1|k-1} = \sum_{j=1}^{\mathcal{J}} \mu^{i|j}_{k-1|k-1} \left[P^j_{k-1|k-1} + \left(\bar{h}^i_{k-1|k-1} - \hat{h}^j_{k-1|k-1} \right)^2 \right] \tag{4.24}$$

The probabilities $\pi^{i,j}_k$ that a transition occurred from state i to state j are calculated according to an FSMC and given by $\pi_{h_i \to h_j,k} = \Pr\{h_{j,k} | h_{i,k-1}\}$, which depends on the statistics of real channel situations as described earlier in Chapter Three.

- Model individual filtering: Individual KFs normally predict and update their state estimates $\hat{h}^i_{k-1|k}$ and covariances $P^j_{k-1|k}$. Innovation sequences and their covariances are also calculated in this part.

- Model probability update: Each KF model probability is updated according to its own innovation sequence. Given the assumption of Gaussian statistics, the likelihood function for the observation can be calculated from the innovation z^i_k and its covariance S^i_k as follows

$$\Delta^i_k = \frac{exp\left(-(1/2)\left(z^i_k\right)^2\left(S^i_k\right)^{-1}\right)}{\sqrt{|2\pi S^i_k|}} \tag{4.25}$$

and the update for predicted model probabilities

$$\mu^i_{k|k} = \frac{\mu^i_{k|k-1}\Delta^i_k}{\sum_{j=1}^{\mathcal{J}} \mu^j_{k|k-1}\Delta^j_k} \tag{4.26}$$

- Combining: The overall CSI estimate $\hat{h}_{k|k}$ and its covariance $P_{k|k}$ are now calculated by combining the weighted individual state estimates $\hat{h}^i_{k|k}$ and covariances $\bar{P}^i_{k|k}$

$$\hat{h}_{k|k} = \sum_{i=1}^{\mathcal{J}} \mu^i_{k|k} \hat{h}^i_{k|k} \tag{4.27}$$

$$P_{k|k} = \sum_{i=1}^{\mathcal{J}} \mu_{k|k}^{i} \left[P_{k|k}^{i} + \left(\hat{h}_{k|k} - \hat{h}_{k|k}^{i} \right)^{2} \right] \tag{4.28}$$

A block diagram of the IMM estimator consisting of $\mathcal{J}$-bank of KFs is illustrated for one cycle in Fig. 4.2. Recall that the IMM-KF estimator does not make a hard decision as to which model is effective at any particular time, but it assigns a probability to each model in the bank set. Typically, the model probability corresponding to the active channel state will be the highest. Based on these mode probabilities, a hard decision as to which model is active can hence be made if so desired.

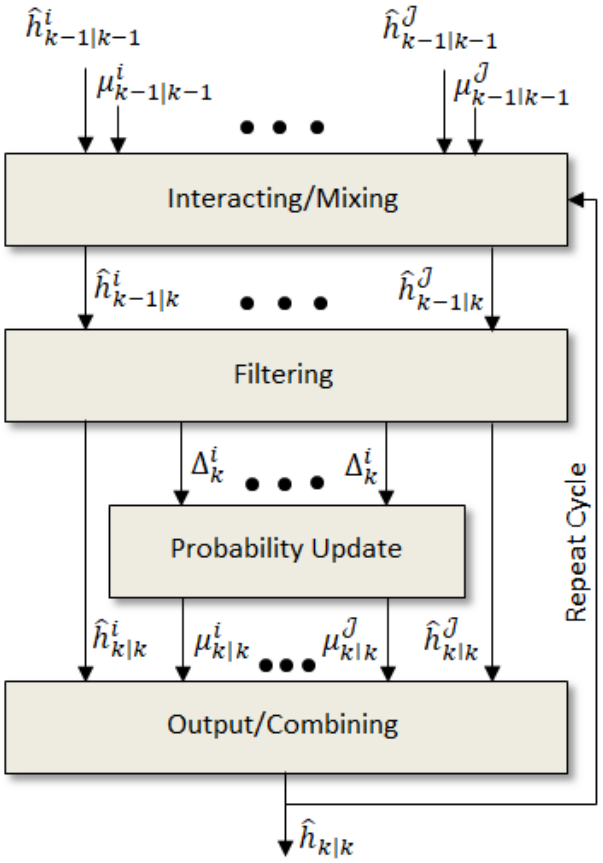


Fig. 4.2 One cycle of the IMM-KF estimator.

4.5 Proposed Adaptive Eigen-Based Estimator

Among countless priority items in a constantly expanding list every once and a while, the speed and complexity of wireless communication systems are of paramount importance. It is always in the best interest of all concerned parties to elevate the speed

to the highest levels possible and reduce complexity as feasible. The MIMO as a key enabling technology can master the first challenge, whereas the SVD method is pronounced as a great champion in reducing signals' subspaces and hence can achieve the second target of less complexity [39, 40, 107, 142]. As was shown in the previous chapter, the SVD is for long considered very competitive in dissolving MIMO channels into their equivalent parallel paths using eigenmode components. These components are virtually highly effective, dominantly influence the overall channels' interactive profile and provide enough information for further processing. Therefore it is relevantly wise to concentrate on such eigenmode significant components instead of handling the whole MIMO transmission scenario, where the latter is an absolute fallacy of excessive system processing time and complexity.

Rank reduction using the SVD method for CSI estimation and subspace tracking is not new [46, 145, 147]; however, it has been often applied using the popular estimation techniques described earlier in the context. The main idea of such classical approaches is to perform the SVD decomposition directly over the channel Hermitian covariance matrix $\boldsymbol{R}_H$ (instead of $\boldsymbol{R}_h$ to cope with MIMO formation), to facilitate the rank reduction and lowering computations. That is, if $\boldsymbol{R}_H = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$, (4.16) can be re-written as [46, 145, 147]

$$\widehat{\boldsymbol{H}}_{MMSE} = \mathbb{E}[\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}] = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}\widehat{\boldsymbol{H}}_{LSE} \tag{4.29}$$

where in this case $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, $\boldsymbol{\Sigma}$ is $diag\{\lambda_1, \lambda_2, \dots, \lambda_{\mathcal{R}}\}$ of the SVs in a descending order, and $\mathcal{R} = min[N_t, N_r]$ denotes the channel rank and also indicates the number of the significant SVs. For WSS process, $\boldsymbol{R}_H$ is liberated from the time index k, so as $\boldsymbol{R}_h$ for SISO. When the SVD is applied over $\boldsymbol{H}$ as was shown in Chapter Three, its entries represent the square-root singular values (SRSV) of the Wishart matrix and is denoted by $diag\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_{\mathcal{R}}}\}$. Whereas applying the SVD over $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ as given in (4.29) involves the diagonal entries of the Wishart matrix eigenvalues.

The aforementioned $\boldsymbol{R}_H$ parameter is generally unknown, but it can be estimated by averaging in time on the receiver side as follows [145]

$$\boldsymbol{R}_H = \frac{1}{K}\sum_{k=1}^{K} \widehat{\boldsymbol{H}}_{LS}(k)\, \widehat{\boldsymbol{H}}_{LS}^{\mathrm{T}}(k) \tag{4.30}$$

The low-rank approximation via SVD looks attractive in reducing the number of multiplications for the channel estimation, however, obtaining the SVD of the covariance matrices is still exhaustive and not economical in practice, especially for the MMSE estimator using OFDM carrier [145]. Therefore, other alternatives need to be identified and here comes another practical paradigm proposed in the work of this study.

Before starting to articulate the new scheme proposed herein, it is advantageous to know about other comparable schemes that are based on eigenmode processing. Without loss of generality, aiming at an effective and straightforward CSI estimate, two powerful yet essentially equivalent simplification strategies can be applied. The first strategy is called as eigenmode transmission (ET), while the second one is the beamforming (BF); both are very popular in the literature [118,119]. The ET is based on SVD and the BF is related to transmit MRC, respectively. Both strategies disintegrate the MIMO channel into $\mathcal{R}$ virtually separated parallel channels and each associated with a gain equals the square root of the i^{th} eigenvalue λ_i of $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ or $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$, where $i \in \{1, 2, \dots, \mathcal{R}\}$.

The application of the BF strategy involves sophisticated array processing [118, 119]. The main structure of which consists of spatial filters devised to receive favourable signals from specific directions and attenuate anything else. It can be equally applied to either one or both of transmission and reception sides. At a large scale, BF refers to signals being transmitted using the principal eigenmodes of a channel [158]. But this is exactly what is meant to be for the ET strategy as a MIMO channel is decoupled into separate and non-interfering ETs. The ET and BF are mainly motivated owing to the treatment of different eigenmodes as scalar channels and thereby preserve the small complexity. In order for both strategies to perform optimally, the provision of full CSI knowledge is stipulated; otherwise, a blind or partial CSI knowledge may yield suboptimal performance, which is also acceptable. In either way, the employment of ET and BF can drive for heightened lifts in the ACM enactment, transmission rates and channel capacities and full diversity can be reached under certain conditions as well [119, 158-160]. The judicious SVD signature on the ET and BF design and processing is substantially phenomenal; it cannot be easily missed or overlooked in any exposed literature in such an active field.

An overwhelming amount of studies endeavoured to use the SVD method, whether for ET or BF, to develop efficiently less-complex channel estimators. To discuss all is a futile attempt; therefore, efforts mostly aligned with the work of this study are only reported herein. Some estimators using the channel covariance matrix have been discussed above, but several others centred on the channel eigenmodes are given special attention in what follows. For example, the possibility of performing signal detection separately on different eigenmodes was indicated in [158], but further elaborations on specific estimation methods were not addressed. The adaptive power allocation for each eigenmode in the feedback link with relevant measuring metric and based on which the ACM scheme can be attended was proposed in [159]. The MLE was implied in that study to build the lower bound performance metric under practical conditions without further estimator details. Moreover, and despite the signal model was explicitly invoked using eigenmodes, the study in [160] opted to employ the non-linear decomposition technique called geometric mean decomposition (GMD). They accordingly came up with the diagonal eigenmodes using the upper triangular portion of the channel covariance matrix disintegrated by SVD, which is commonly named quadrature right decomposition (QRD), and the estimation was done by applying the successive interference cancellation (SIC) algorithm.

More into cross-referencing, a number of studies typified on the explicit eigenmodes inclusion in the signal model can be pulled out from the literature [160-165]. A unified treatment of meriting the straight action of eigenmodes estimation remains potentially very slim among these and other unlisted attempts. For example, signal models encompassing eigenmodes for the purpose of evaluating the overall transmission aspects can be readily recognised in [161-163]. A derivation framework for the statistical distribution of various eigenvalues and associated Wishart matrix was proposed in [161]. The average BER and outage probability performances of spatial multiplexing (SM) MIMO systems in Rayleigh channels were approached in [162], while the performance of MIMO systems employing multichannel beamforming in arbitrary-rank Rician channels was examined in [162]. The same eigen-based signal modelling approach is adopted in all of these studies, and also in [164, 165]. Slightly different from the above, the ML and LMMSE algorithms were directly applied on the eigenmodes to estimate their values in [164, 165]. All these recent studies shared a

common outcome of complexity reduction and enhanced performance due to the SVD involvement as a vivid tool to produce channels with lesser ranks and subspaces.

Inspired by the above, the substantiation of the new paradigm proposed in the work of this study becomes quite apparent. The proposed approach here departs from other roadmaps suggested earlier in the sense that it is not concerned with the factorisation of the channel covariance matrix $\boldsymbol{R}_H$. The core idea of the proposed paradigm is mainly centred on the eigenvalues of the decomposed channel, and instead of being treated individually they are summoned to generate a new single parameter. This parameter symbolises the variant effects of a block-fading channel that is readily available to streamline the IMM-KF estimation algorithm. Such an approach is by far believed to be unique and there is no evidence that it has been experimented elsewhere previously.

Now, consider a flat-fading channel is stimulated by an $N_t \times N_r$ MIMO scheme, where N_t stands for transmission antennas and N_r stands for receive antennas, respectively. The signal model and channel coefficients matrix $\boldsymbol{H}$ were described earlier in the context as per (3.3) and (3.4) respectively. By applying the SVD procedure, the signal model is simplified in terms of the eigenmodes equivalent representation (3.24), or (3.25) for STBC, and repeated down here for the purpose of exposition completion and $\forall k \in K$

$$y(k) = \sum_{i=1}^{\mathcal{R}} \sqrt{\lambda_i}\, x(k) + n(k) \tag{4.31}$$

where $i = 1, 2, \ldots, \mathcal{R}$ and λ_i denotes the channel eigenmodes and also the eigenvalue entries of the Wishart matrix, $\mathcal{R}$ is the channel rank and n is the AWGN as both defined earlier. It can be easily deduced that the above eigen-based signal model matches those given in [160, 161] and [163, 164], while the models of [162, 165] are slightly different despite sharing the same meanings and interpretations, is a matter of flexible representation only.

As stated earlier, the intention here is to summon the above given eigenmodes in order to come up with a one single parameter reflecting on the channel variations and also to contribute to achieve more computational simplification. The newly generated parametric is called effective channel, which represents to summation of all active

eigenmodes. It can be taken as $\hbar = sum.diag\left[\sqrt{\lambda_{\mathrm{i}}}\right]$ and deployed by replacing h in (4.31) as follows

$$y(k) = \hbar x(k) + n(k) \tag{4.31}$$

where the time index, k, is not attached to $\hbar$ as the latter remains constant over the observation processing time, i.e., the channel is of a block or flat fading effect. To build the state-space model suitable for KF application, the model given earlier in (4.17) is to be revised. Hence, the following AR(1) Gauss-Markov model can embrace most of the channel dynamics and lead to effective tracking of the new eigenmode parameter

$$\hbar(k) = \hbar(k-l) + w(k) \tag{4.32}$$

where the statistics of $w(.)$ process remain unchanged, which is attributed to the law of large numbers and central limit theory (CLT) as the sum of many processes, whether of the same nature or not, converges almost surely to the statistics of a white Gaussian noise [107, 142, 143]. The Wishart matrix statistics also remain unaltered unless massive antennas involved as per the same aforesaid theories. The new concept of total square root of effective eigenmodes, which is mathematically equivalent to the Frobenius norm, will be dubbed Frobenius ET (FET) in this study. Led by the styles of SVD and parallel eigen paths in [39, 160], the simplified adaptive IMM-KF estimator is portrayed in Fig. 4.3 below.

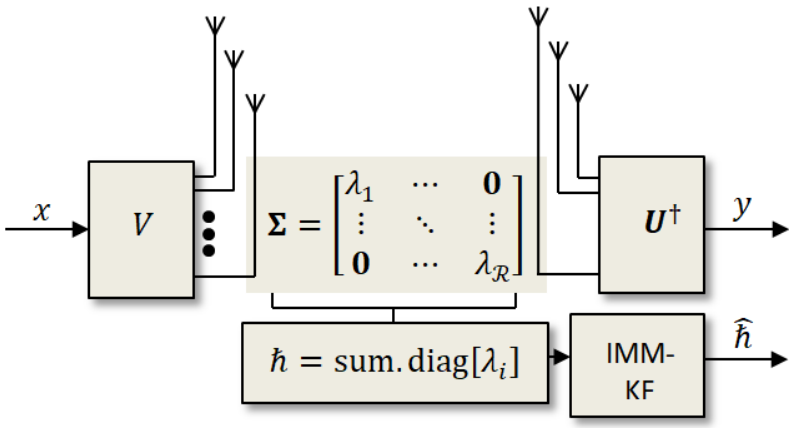


Fig. 4.3 Simplified diagram of the proposed eigen-based adaptive estimator.

Ensuing to the analysis of new FET paradigm, the corresponding $\mathcal{J}$-bank algorithms for IMM-KF adaptive estimator are recompiled and summarised as in Table 4.1 below.

Table 4.1 IMM-KF adaptive estimator for FET decomposition.

Item	Algorithms	Notes
KF	$\begin{cases}\widehat{\hbar}_{k\|k-1}=\widehat{\hbar}_{k-1\|k-1}\\ P_{k\|k-1}=P_{k-1\|k-1}+\sigma_w^2\end{cases}$	KF gain G_k, Residual sequence $z_k=\left(y_k-\widehat{\hbar}_{k\|k-1}\right)$ of covariance $S_k=P_{k-1\|k-1}+\sigma_w^2$
	$\begin{cases}\widehat{\hbar}_{k\|k}=\widehat{\hbar}_{k\|k-1}+G_k z_k\\ G_k\triangleq P_{k\|k-1}{S_k}^{-1}\\ P_{k\|k}=(1-G_k)P_{k\|k-1}\end{cases}$	
IMM-KF	$\mu_{k-1\|k-1}^{i\|j}=\frac{\pi_k^{i,j}\mu_{k-1\|k-1}^{i}}{\mu_{k\|k-1}^{j}};\ \mu_{k-1\|k-1}^{i}=\sum_{j=1}^{\mathcal{J}}\pi_k^{i,j}\mu_{k-1\|k-1}^{j}$	Interact, then normal individual filtering
	$\overline{\hbar}_{k-1\|k-1}^{i}=\sum_{j=1}^{\mathcal{J}}\mu_{k-1\|k-1}^{i\|j}\widehat{\hbar}_{k-1\|k-1}^{i};$	
	$\bar{P}_{k-1\|k-1}^{i}=\sum_{j=1}^{\mathcal{J}}\mu_{k-1\|k-1}^{i\|j}\left[P_{k-1\|k-1}^{j}+\left(\overline{\hbar}_{k-1\|k-1}^{i}-\widehat{\hbar}_{k-1\|k-1}^{j}\right)^2\right]$	
	$\Delta_k^i=\frac{\exp\left(-(1/2)\left(z_k^i\right)^2\left(S_k^i\right)^{-1}\right)}{\sqrt{\lvert 2\pi S_k^i\rvert}};\ \mu_{k\|k}^{i}=\frac{\mu_{k\|k-1}^{i}\Delta_k^i}{\sum_{j=1}^{\mathcal{J}}\mu_{k\|k-1}^{j}\Delta_k^j}$	Update
	$\widehat{\hbar}_{k\|k}=\sum_{i=1}^{\mathcal{J}}\mu_{k\|k}^{i}\widehat{\hbar}_{k\|k}^{i}$	Combine
	$P_{k\|k}=\sum_{i=1}^{\mathcal{J}}\mu_{k\|k}^{i}\left[P_{k\|k}^{i}+\left(\widehat{\hbar}_{k\|k}-\widehat{\hbar}_{k\|k}^{i}\right)^2\right]$	

Not startling, the FET concept can also be applied to classical estimators to bonus extra complexity discounts. As per (4.10), the LSE and MLE remain intact under uncorrelated white Gaussian noise, while the MMSE estimator in (4.16) can easily be reduced to

$$\hat{h}_{MMSE}=\sigma_{\hbar}^2(\sigma_{\hbar}^2+1/\gamma_s)^{-1}\hat{h}_{LS} \tag{4.32}$$

where $\boldsymbol{R}_h$ is now reduced to $\sigma_{\hbar}^2$, thanks to the low-rank subspace achieved due to the SVD powerful touch. Relevant complexity savings are addressed in the next section.

4.6. Complexity Analysis

The complexity analysis is worked out here to assess the benefits of different estimation algorithms. A comparison between the computational complexities of the

proposed IMM-KF estimator and other discussed estimators using the new FET decomposition is engaged in this part. It is known that the complexity and computational analysis hinges on interrogating the numbers of complex additions and multiplications. Rather than perfecting numbers, it is usually reasonable that these numbers are roughly calculated. To that note, there are no unique figures that can be given in relation to the exact computational operations requirement for any algorithm; it is mainly dependent on how approximations and assumptions are made.

Table 4.2 Total numbers of dominant complexity or computational operations.

Item	MIMO	FET (SVD)	Notes
KF	$\mathcal{O}(\mathcal{R}^3)$	-	FET (SVD) almost trivial
IMM	$\mathcal{O}(\mathcal{J}^2)$	$\mathcal{O}(\mathcal{J}^2)$	
IMM-KF	$\mathcal{O}(\mathcal{J}^2\mathcal{R}^3)$	$\mathcal{O}(\mathcal{J}^2)$	
LS	$\mathcal{O}(\mathcal{R})$	-	FET (SVD) almost trivial
ML	$\mathcal{O}(\mathcal{R}^2)$	-	FET (SVD) almost trivial
MMSE	$\mathcal{O}(\mathcal{R}^3)$	-	FET (SVD) almost trivial

To leverage more flexibility, it is assumed that the complexity analysis is performed over one time sample and $\mathcal{R} = N_t = N_r$ for a full-rank singular-value matrix $\boldsymbol{\Sigma}$. The SVD is embedded in the FET approach and they both are assumed to have the same computations. For BPSK signals, the summation process of effective eigenmodes is trivial and therefore not counted in the overall calculations intake. By examining the structures of the above given estimators, the final figures of complexity and computational operations can be easily deduced as shown in Table 4.2. As can be seen in this table, the FET with embedded SVD algorithm makes a great amount of complexity reduction almost equivalent to the case of having scalar variables being submitted for estimation. The MIMO case without rank reduction is noticeably the most demanding of all estimators. That can be mainly attributed to the heavy operations involved in matrix inversions, which are assumed to have the same complexity of Gauss-Jordan elimination process of $\mathcal{O}(\mathcal{R}^3)$ additions and $\mathcal{O}(\mathcal{R}^3)$ multiplications [40]. Insights on these computations can be found in [40, 46, 149, 166-168], which were followed here but manipulated to yield the dominant figures only.

4.7. Simulation Results and Discussion

The performance corroboration of the proposed adaptive IMM-KF estimator and traditional estimators is offered using simulation exercises in this section. The innovative paradigm of FET subspace reduction is incorporated in the signal model for all estimators. Two transmission scenarios are considered; the first comprises a 2×2 MIMO while the second is of 3×3 MIMO. The signalling is evoked by a BPSK sequence of 129 samples length and unity power sent in the direction of a Rayleigh channel. The channel is assumed with a plausible flat fading over certain block samples. An AWGN is also incurred in these channels and the SNR is 10 dB. In each simulation scenario, the 129 samples are divided into three block segments; the length of each is 43 samples, to manifest fixed fading. The segmented channel fading magnitudes along their corresponding SRSVs are given in Table 4.3 below. The three segments of the first channel exhibit severe, medium and mild fading behaviours, respectively, while the second channel is assumed to have the opposite behaviour. These arbitrary fading attributes are plotted in Fig. 4.4 and Fig. 4.5 below.

Table 4.3 Scenarios of MIMO channel simulation settings.

Item	Scenario 1 (2×2)	Scenario 2 (3×3)
$\boldsymbol{H}$ **blocks**	$[0.1\ \ j0.1;\ j0.1\ \ 0.1]$ $[0.5\ \ j0.5;\ j0.5\ \ 0.5]$ $[0.9\ \ j0.9;\ j0.9\ \ 0.9]$	$[0.9\ \ 0.9\ j0.9;\ 0.9\ \ j0.9\ 0.9;\ j0.9, 0.9, 0.9]$ $[0.5\ \ 0.5\ j0.5;\ 0.5\ \ j0.5\ 0.5;\ j0.5\ 0.5\ 0.5]$ $[0.1\ \ 0.1\ j0.1;\ 0.1\ \ j0.1\ 0.1;\ j0.1\ 0.1\ 0.1]$
$\sqrt{\boldsymbol{\lambda}}$ **SRSV**	$[0.14\ \ 0.14]$ $[0.7\ \ 0.7]$ $[1.27\ \ 1.27]$	$[2\ \ 1.27\ 1.27]$ $[1.11\ \ 0.7\ 0.7]$ $[0.22\ \ 0.14\ 0.14]$
FET	0.28 1.4 2.54	4.54 2.51 0.5

From these illustrations, it is quite evident that each single path of decomposed channels carries important chunks of the overall transmission power and hence it is healthier to be summoned rather than being treated individually or merely use the dominant maximum eigenmode for transmission only, which is the case of ET. This is the main idea behind the meaningful postulation of the new paradigm of FET proposed here. Such an approach consequently keeps as much channel power as possible and

maintains an efficient track of its varying conditions instead of being aborted and vanished in nowhere. The power components of different channel paths have a direct impact on the transmitted signal power and hence it is more advisable to be all consolidated and exploited.

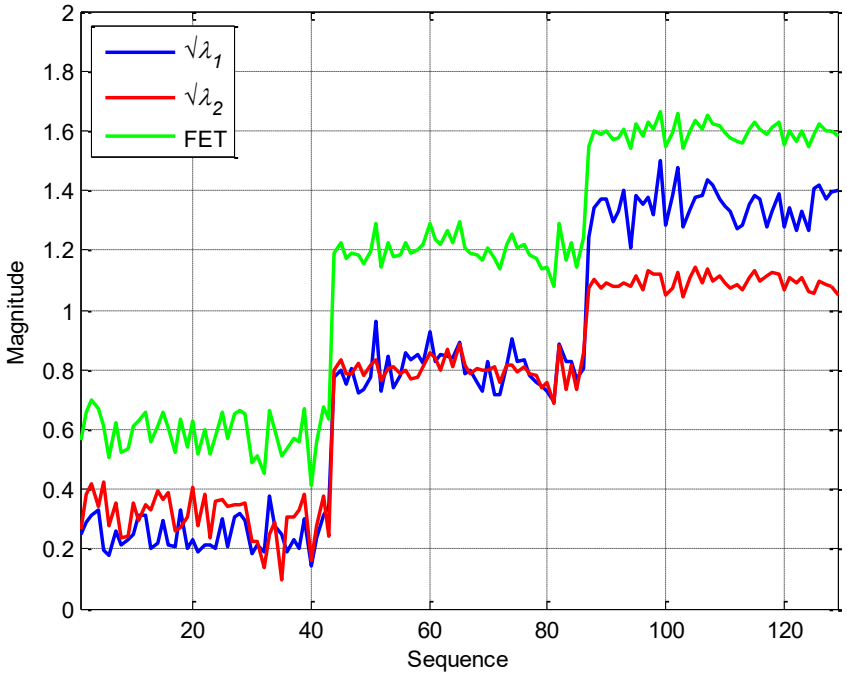


Fig. 4.4 SRSVs of 2 × 2 block-fading Rayleigh channel in AWGN.

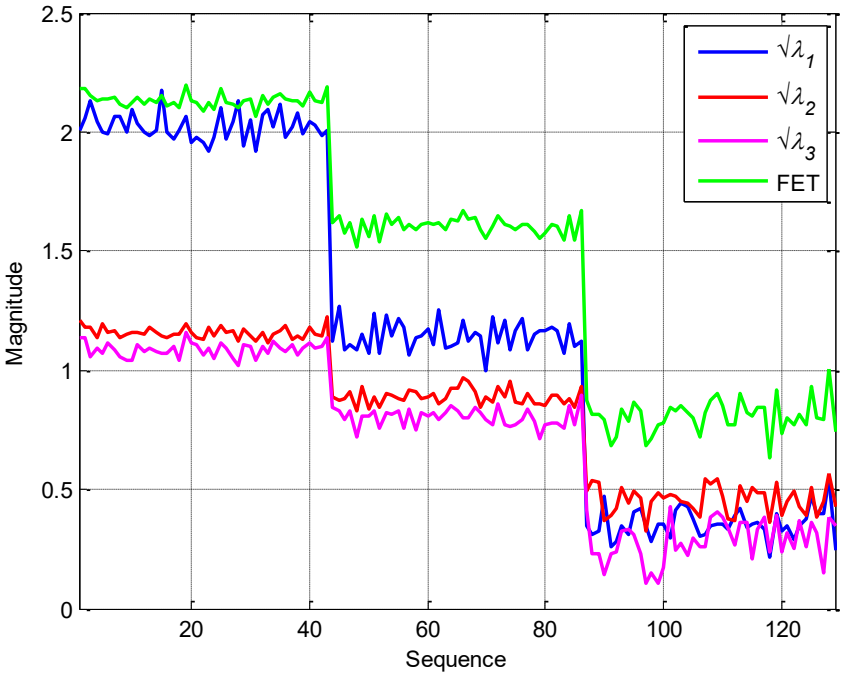


Fig. 4.5 SRSVs of 3 × 3 block-fading Rayleigh channel in AWGN.

Among the conventional channel estimators described earlier in the context, two only are assessed; namely, the LSE and MMSE. That is because the MLE and MAPE are equivalent to the LSE and MMSE under uncorrelated white Gaussian noise impairments. Therefore, a comparison between the LSE and MMSE performances would be sufficient to gain a picture of the rest. The performances of the proposed KF and IMM-KF algorithms are also validated against the abovementioned, where the initialisations of which are arbitrated as given in Table 4.4 below. A bank of $\mathcal{J} = 3$ KFs is considered for the IMM algorithm whether over 2×2 or 3×3 MIMO transmissions. On the other hand, there is no parametric initialisation necessary for the LSE and MMSE schemes.

Table 4.4 Initialisation of KF and IMM-KF algorithms.

Item	Parameters Initialisation	Notes
KF	$\hat{h}_{0\|0} > 0, P_{1\|0} \geq 1$ $\sigma_n^2 = 0.01, \sigma_w^2 = 0.1$	Same for each KF in the IMM except process noise
IMM-KF	$\sigma_{w1}^2 = 0.1, \sigma_{w2}^2 = 1, \sigma_{w3}^2 = 5$ $\mu_{0\|0}^i = [0.8\ 0.1\ 0.1], \forall i \in [1,2,3]$ $\pi_0^{i,j} = \begin{bmatrix} 08 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}, \forall i,j \in [1,2,3]$	Individual process strengths Modal probabilities Transitional probabilities

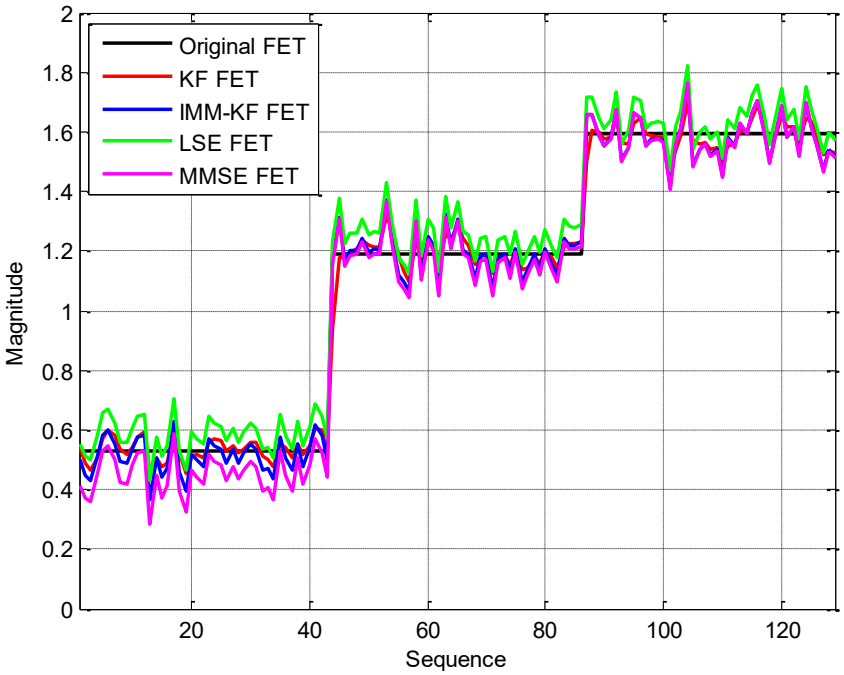


Fig. 4.6 Estimation of SRSVs over 2×2 block-fading Rayleigh channel in AWGN.

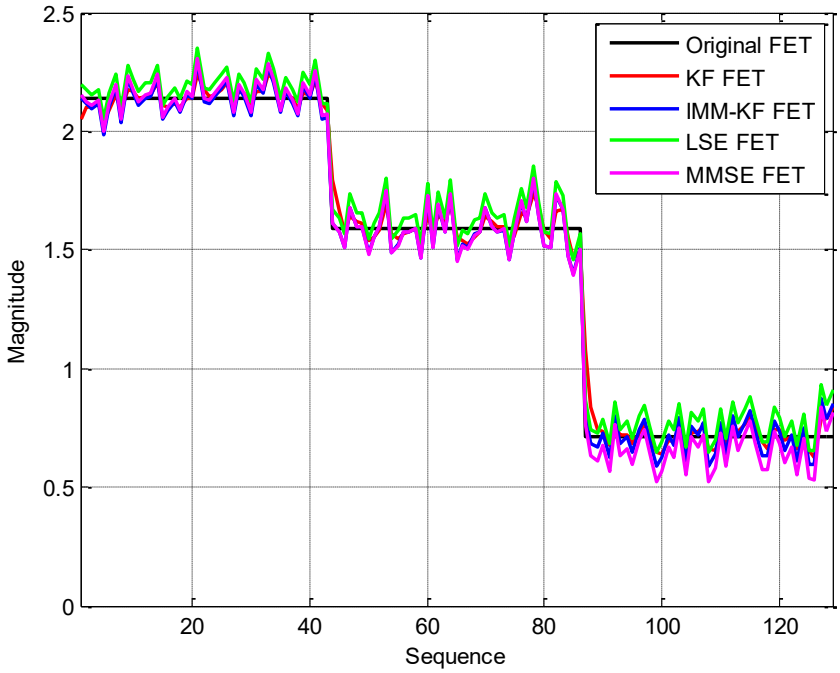


Fig. 4.7 Estimation of SRSVs over 3 × 3 block-fading Rayleigh channel in AWGN.

The estimators are separately allocated through the given channels and the simulation outcomes are compiled in Fig. 4.6 and Fig. 4.7 for each MIMO scenario, respectively. Both figures show that all estimators maintain a good track of the channels' SRSV variations; yet some perform better than others. The KF and IMM-KF obviously outperform the LSE and MMSE and among which the IMM-KF has more preference especially during the channels' deep fading parts. The IMM-KF is also superior through the channels' transitions between states, which is inherited from the FSMC's natural behaviour.

The main driver behind such robust adaptation performance of the IMM-KF is centred on the attractive switching competence of the IMM structure itself. If the channel has mild fading, the KF with the lowest state noise power $\sigma_{w1}^2 = 0.1$ is assigned the highest weight in the output of probability combining. As the fade moderates, the state noise power $\sigma_{w2}^2 = 1$ is allocated the next highest weight than other KFs in the IMM bank. The deep fading effect is treated with the highest state noise power $\sigma_{w3}^2 = 5$ and the output of combining probability gets the highest weight accordingly. The first encounter is similar to the situation of having a lone KF with action virtually reminiscent of a low-pass filter (LPF). This LPF reduces the channels' harsh fluctuations and smooths out the output as revealed in the above two plots.

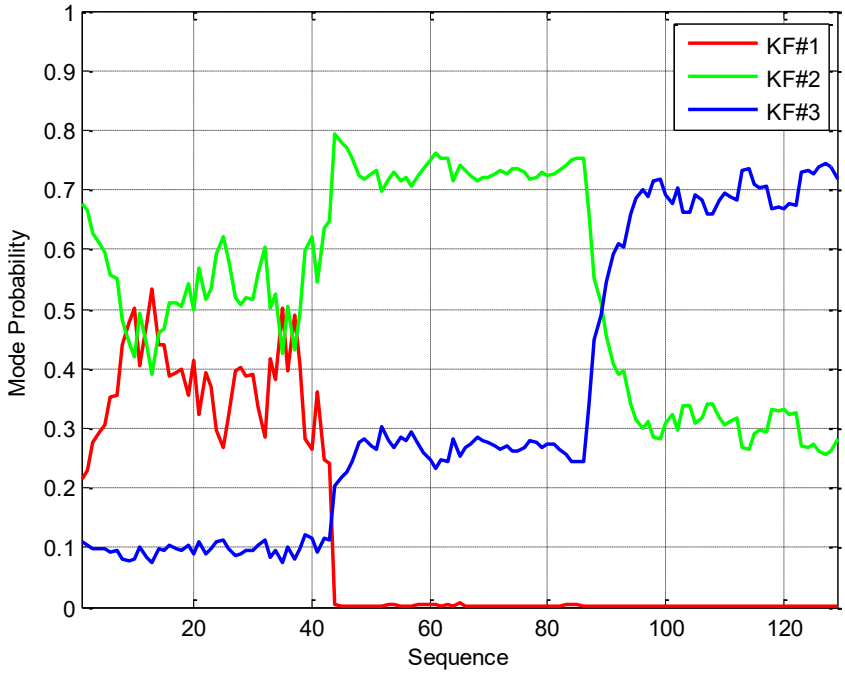


Fig. 4.8 Modal probabilities of 3-bank KFs over 2 × 2 MIMO block-fading channel.

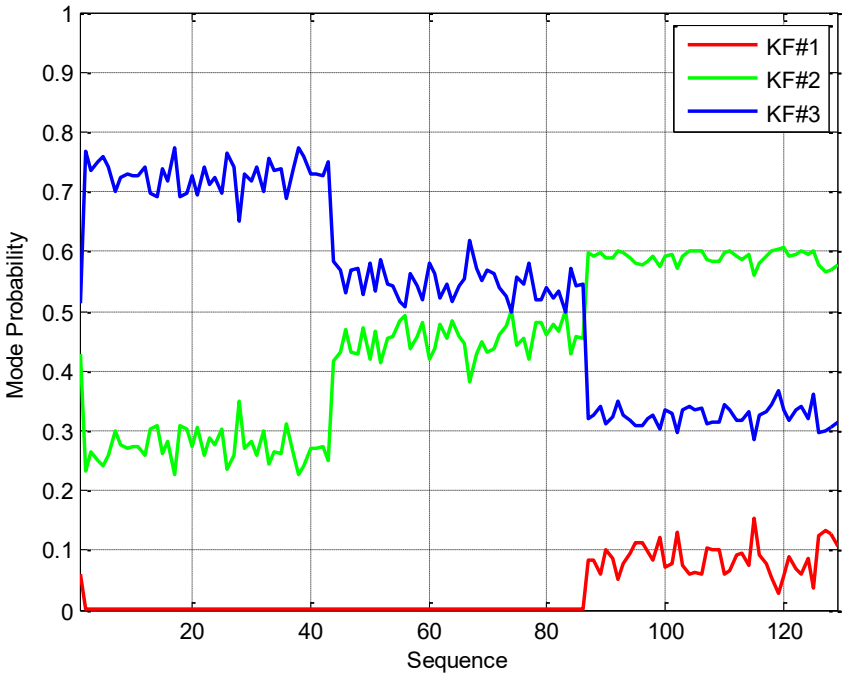


Fig. 4.9 Modal probabilities of 3-bank KFs over 3 × 3 MIMO block-fading channel.

Moreover, the transitional probabilities of 0.8 and 0.1 represent the percentage of time a state spends about 80% in its current positions and makes a transition to adjacent cells, whether forward or backward, with 10% likelihood occurrence of each. However, the modal and transitional probabilities change with respect to the innovation behaviour of

each KF branch as the time elapses. The perspectives of such IMM-KF switching probabilities to cope with the various jumping fades, are well discerned in Fig. 4.8 and Fig. 4.9 for each simulation scenario, respectively.

To furtherly validate between the few estimators selected here, the last exercise is focused around the design of an adequate measurement metric commonly employed in parametric applications. The performance of different adaptive estimators can be generally interrogated by applying the following root-mean square error (RMSE) formulae over all K sequence samples [168]

$$RMSE_{\hbar}(K) = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\widehat{\hbar}(k) - \hbar(k)\right)^2} \tag{4.33}$$

which can be easily transformed into the following recursive expression

$$RMSE_{\hbar}(k) = RMSE_{\hbar}(k-1) + \sqrt{\left(\widehat{\hbar}(k) - \hbar(k)\right)^2 (k-1)/k} \tag{4.34}$$

For the purpose of simplicity of illustration, only 43 samples of one block fading are used in the simulations. The RMSE of all estimators are computed over both 2×2 and 3×3 MIMO channels and the outcomes are rendered in Fig. 4.10 and Fig. 4.11, respectively.

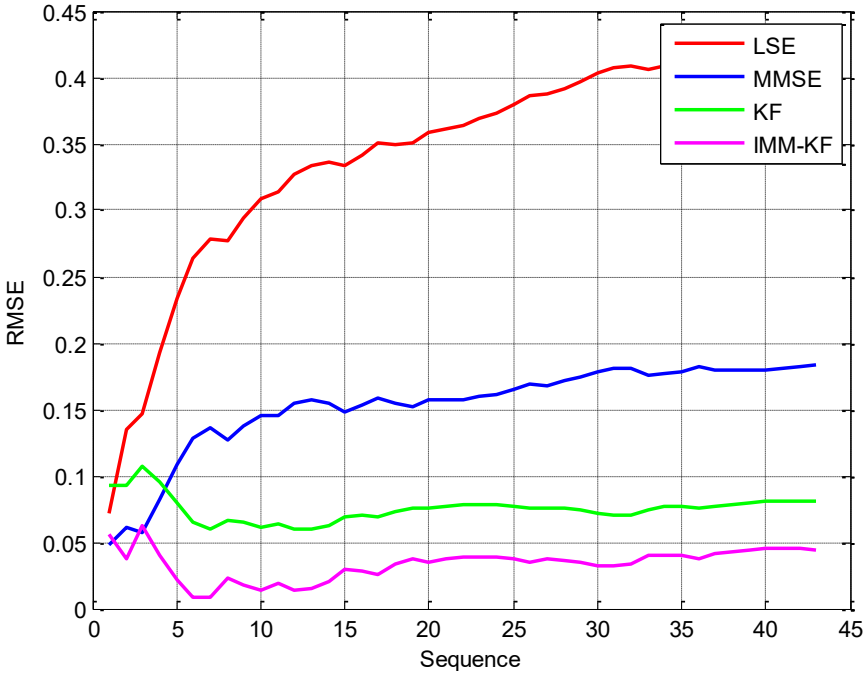


Fig. 4.10 RMSE comparison of channel estimators over 2×2 fading channel.

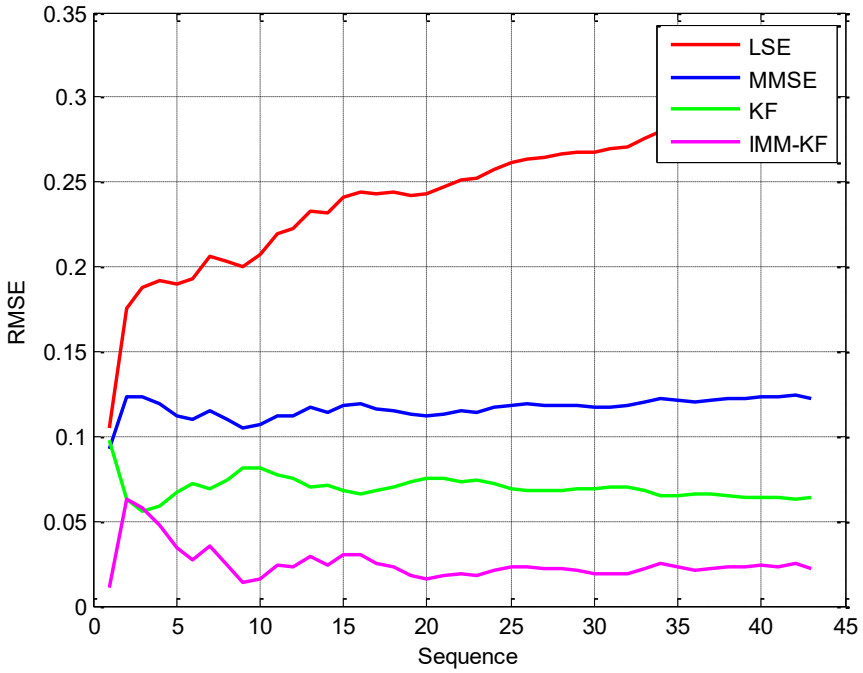


Fig. 4.11 RMSE comparison of channel estimators over 3 × 3 fading channel.

The viability of the KF and IMM-KF channel estimators keeps pitching high. The RMSE results leave no doubt about their superior performances compared to other estimators, and the IMM-KF is clearly healthier than that of KF. Despite the incremental trend of the RMSE formulae, the KF and IMM-KF nearly maintain the errors within very low limited margins, while the MMSE barely holds the same. On the other hand, the LSE clearly has the worst trend among others as the error of which keeps shooting high monotonously. Many simulation exercises were applied over other fading segments or continuously changing fade, but there is not enough space to include all.

4.8 Summary

This chapter has laid down the foundations of estimation theory that is useful in the business of wireless communication systems, specifically in terms of channel tracking and estimation applications. The classical channel estimators (LSE, MLE, MMSE, MAPE) were illustrated and compared to a proposed new paradigm using analysis and simulation methods. Namely, the deployment of eigen-based IMM-KF is substantiated to be a very appealing adaptive estimator that can cope with a variety of challenging

work conditions. The subspace reduction of MIMO channels was illustrated using a newly suggested approach called FET, which has the SVD embedded in its structure. The main observation that can be concluded from the main outcomes of this chapter is that the FET and SVD algorithms are strikingly powerful in saving complexities and computations and hence they will continue to play a lead role in massive MIMO, MU and multi-carrier like OFDM for many years to come. Without loss of generality, all results also agreed to the above findings and affirmed the IMM-KF based on FET as a superior adaptive channel estimator among all others tackled in this study. It is undoubtedly forecasted to continue curbing the highways of diversified and resourceful adaptation techniques, and therefore will be re-enacted yet again in the newly suggested ACM, AMC and SS paradigms as will be shortly seen in the subsequent chapters.

Chapter 5

ADAPTIVE CODED MODULATION

5.1 Introduction

This chapter addresses the main ACM techniques that have been prevailing in most wireless communication systems for a long time and are expected to survive for many years to come. Among which, the signal power, modulation pattern and rate, and channel coding techniques constitute the main ingredients in wireless communication systems of high quality. The success of such techniques in achieving their designated goals is largely determined by the provision of a feedback channel to furnish the CSI status for transmitters to make appropriate adjustments. The setting of ACM techniques is hence expected to be of duplex type rather than simplex in most wireless communication systems, which means exchanging information in both directions. To fulfil the ACM objectives, effective WF procedure and switching strategies between constellation headers are employed. The common BICM and TCM, as efficient coding methods to combat channel fading, are also introduced. Whereas the main contribution of this chapter is signified by exploiting the IMM-KF algorithm to attend the estimation of various CSI parameters and feed them back for transmitters and upon which to decide the proper course of adaptation actions. Namely, the IMM-KF is applied to estimate the noise power and SNR values as well as the channel fading coefficients in order to enable the desirable ACM functionality.

5.2 Adaptation Techniques

ACM techniques are broadly meant to improve the spectral efficiency of radio links for predefined levels of maximum QoS or minimum BER. The notion of ACM dealing with various channel conditions is not new and can be traced back to the 1970s. However, it was not widely spread until much later in the middle of the 1990s when many serious studies started to thrive upon the discovery of optimum adaptation structures [39, 40, 43].

In a wide sense, wireless communication systems with fixed modulation are not competent with concern to the capacity and quality of fading channels [39, 40, 43]. To achieve robust and spectrally efficient communication over multipath fading channels, ACM techniques exploiting the CSIT are highly recommended. Unlike non-adaptive schemes, which are designed to achieve acceptable performance for worst-case channel scenarios, the ACM takes the advantage of favourable channel conditions by allocating power, rate and coding efficiently. The CSIT conversed through reliable feedback link is a key for better control of the signal conditioning with respect to unforeseen channel disorders. Referring to Fig. 3.1 and to maintain consistency with the ACM representations of [39, 40, 43, 44, 47-49, 51, 53], a simplified ACM diagram focusing on the essential components can be devised as depicted in Fig. 5.1. SISO is shown in Fig. 5.1, which can be easily rearranged for a MIMO scheme.

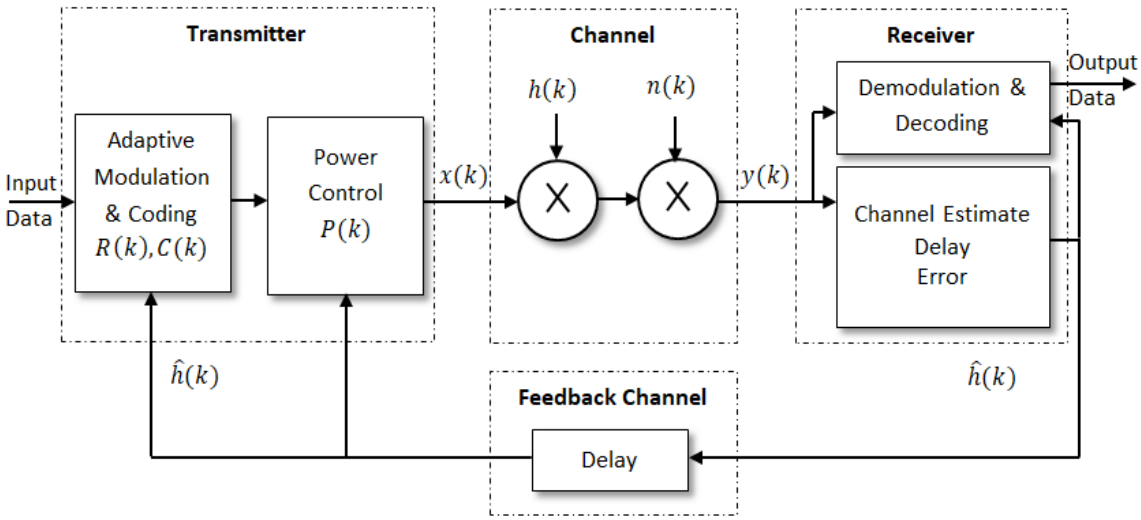


Fig. 5.1 Generic ACM system model.

It is apparent from Fig. 5.1 that the estimated parameters of the forward channel, or downlink, returned via feedback channel, or uplink, are in the essence build-up of any functional ACM system. The CSIT and CSIR can be reciprocally exchanged by transmitter and receiver (in short "transceiver") in both directions [40, 44]. This forms closed-loop signalling to achieve reliable adaptation. The transceiver in this case can adapt its signalling by using the information acquired during the receiving mode. Most central, the engagement of a return link should always be emphasised in order to guarantee a flawless ACM functionality, without which no such thing can be invoked at all.

In ACM, the modulation order depends on the channel reliability. When the channel has high reliability, the modulation order is increased to maximise the spectral efficiency and throughput; when the transmission channel is unreliable, the modulation order is decreased to provide acceptable BER. Elementary coded modulation (CM) techniques designed for AWGN channels can be abstracted in very generic ACM forms and produce coding gain in addition to modulation gain. The size, power and symbol time of transmitted signal constellation can hence vary to maximise the average data rate without affecting coding gain. Due to the attractive features of multiple ACM recipes, immense research endeavours have been consequently induced, especially in the area of emerging CR applications.

All the aforementioned adaptation schemes generally aim at maximising the average transmission rate subject to adhering to BER specifics. Intuitively, the BER depends on the inner distances of constellation structures scaled by the channel gain. If the channel gain is large, transmitting a large constellation is affordable. If the channel gain is faded, the size of the constellation needs to be reduced in order to guarantee the same BER target. Integrating into the real world of wireless communication systems' standards, the IEEE 802.11x and IEEE 802.16.x family of standards for WLANs and WiMAX, respectively, set good examples for high data rate in cellular data communications and wireless/mobile internet access systems [44].

Generally, the selection of an ACM technique is highly determined by the best possible use of resources available for transmission, i.e., bandwidth, power and complexity, in order to achieve the targeted QoS. The latter is typically expressed in terms of BER, which in turn has a decreasing trend with respect to SNR values. For the

design of ACM to be sensible, the following four strategies need to be carefully examined [39, 40, 42]:

- BER: indicates system reliability.
- Bandwidth efficiency: measures efficiency in bandwidth overheads.
- SNR: measures how the available power is efficiently used for certain QoS levels.
- Complexity: measures the cost of equipment.

Whether applied to wire or wireless channels, the majority of the ACM techniques can be categorised into the following basic groups [39, 40, 42, 44, 169]:

- Adapting Power Level: All transmission parameters are unchanged when the channel varies, but the transmission level is varied according to the channel fluctuations through power control. This approach increases the peak power necessary for transmission. If the channel is accessed by more than one user, the amount of interference power can also be controlled.
- Adapting Constellation Size: Instead of manipulating the power, the adaptive modulation increases the data transmission efficiency by increasing the size of a hierarchy of different constellations. While keeping the channel state continuously monitored, the modulation index and symbol rate can be varied to adapt the modulation scheme to the channel SNR values.
- Adapting Code Rate: In response to the channels variations, the rate of particular coding schemes can be altered accordingly. Among various coding schemes, the punctured convolutional codes are especially useful to satisfy this purpose. That is because they enable encoding and decoding processes to be performed adaptively while keeping the encoder and decoder basic structures intact.
- Adapting Power Level and Constellation Size: Both modulation scheme and transmit power level can be adapted as per the channel variations in a single or multiple users' environment. This adaption merger results in significant throughput boosts compared to the situation without power control.
- Adapting Constellation Size and Transmission Rate: Both constellation size and symbol rate can be adapted in this approach with respect to the channel

variations. The system in this case selects the optimum modulation parameters in order to maximise the bit rate while maintaining the required BER. A lower symbol rate can be achieved by consecutively transmitting identical modulation symbols at the maximum symbol rate, which is equivalent to the repetition coding in this case.

- Adapting Power and Transmission Rate: In order to maximise the spectral efficiency while satisfying average power and BER constraints, both the transmission rate and the power level can be adaptively controlled accordingly.

- Adapting Modulation Size and Coding, or TCM: By fixing the encoder information bits, and adapting the uncoded bits as per the CSI estimate attributes, the trellis structure is changed. Such adaptation, however, is not robust against estimation errors, especially under Rayleigh fading channels. Owing to its increasing the time diversity, the BICM may hence be more attractive instead. The code is kept fixed in BICM, while the constellation is adapted to the channel conditions (hence, it can be categorised under the header "adapting the constellation size"). Analyses suggest that the code structure of BICM is more suitable for adaptive systems to support users of high mobility.

- Adapting Code Rate, Symbol Rate, Constellation Size: All these adaptation attributes can be performed simultaneously. Code rate adaptation is obtained by puncturing a convolutional code, while the constellation size is selected by setting SNR thresholds. As in any generic case, the system transmits no data if the target BER cannot be achieved under any combination of parameters.

5.3 Adaptive Rate and Power

The adaptation by varying transmission constellation pattern and power has an abreast of being well recognised as a radical choice in reality and hence endeavoured in this study [39, 40, 43, 47, 49, 51, 53]. The attention first is to limit the link adaptation to uncoded MQAM signalling immersed in AWGN without fading. Suppose each controllable parameter stands for one DoF. With no coding, the two DoFs addressed here are the choice of modulation formats in different SNR regions, and the possibility of varying the transmit power within these regions [48, 51]. Generally, studies have

shown that using just one or two DoFs in adaptive modulation yields close to the maximum possible spectral efficiency obtained by utilizing all DoFs. Therefore, the choice of parameters to adapt is usually constrained by implementation considerations [51]. Before embarking onto further analysis, it is worth to recall that constellations can be denoted by different sizes M, which correspond to different symbol rates R. To avoid probable misconception, the notion of R is used here instead of $\mathcal{R}$, where the latter is more associated with spatiotemporal coding rate as explained earlier in Chapter Three. Recall that the calculation of $\mathcal{R}$ can be explicitly interpreted in terms of M and R, in addition to other spatiotemporal factors.

Few parameters are due for reckoning and in what follows, like previous and what follows, is served by linear modulations, such as MQAM. In such a case, the adaptation takes place at a multiple of the symbol rate $R_s = 1/T_s$, and accordingly the system has a Nyquist bandwidth $B = 1/T_s$. The channel is flat-fading and modelled using discrete sequence of T_s symbol time. The channel has stationary and ergodic time varying coefficient $h(k)$ that follows a given PDF $p(k)$ and $n(k)$ is the AWGN with two-sided PSD $N_0/2$. The instantaneous received SNR is given by $\gamma_s(k) = \bar{P}h(k)/N_0B$, where $\bar{P}$ is the average transmit signal power, $0 \leq \gamma_s(k) < \infty$ and its expected value over all time is $\bar{\gamma}_s = \bar{P}\bar{h}/N_0B$. Recall that this SNR is equivalent to the expressions (3.6) and (3.7) for ergodic processes given earlier in the context. Since $h(k)$ is stationary, the distribution of $\gamma_s(k)$ is independent of k and the PDF of which can hence be denoted as $p(\gamma_s)$. Furthermore and for simplicity of exposition, both the forward and backward channels are assumed ideal and without delay.

The estimated power gain or the received SNR at time k is used to adapt the modulation parameters at the transmitter, namely; data rate, constellation, and transmit power, i.e., $R_s(k), C(k)$ and $P(k)$, respectively. Recall from Chapter Three, but now in an instantaneous form, for M-ary modulation the data rate is given by $R_s(k) = \log_2 M(k)/T_s = B\log_2 M(k)$ bps. The spectral efficiency of the M-ary modulation is $R_s(k)/B = \log_2 M(k)$ bps/Hz. The estimated value of SNR is $\hat{\gamma}(k) = \bar{P}\hat{h}(k)/N_0B$, which is based on the power gain estimate $\hat{h}(i)$. The adaptation parameters can then be denoted by $R_s(k) = R_s(\hat{\gamma}(k))$, $C(k) = C(\hat{\gamma}(k))$, and $P(k) = P(\hat{\gamma}(k))$, and the time parameter can be omitted for simplicity to yield $\hat{\gamma}$, $R_s(\hat{\gamma})$, $C(\hat{\gamma})$ and $P(\hat{\gamma})$. It is assumed that the estimate $\hat{h}$ of the channel power gain h is available to the receiver and returned

to the transmitter without delay. This is an ideal situation and merely intended to simplify the presentation.

5.3.1 BER Upper Bounds

In the premise of real-time adaptation, there is a delicate balance of the link budget through varying the transmitted power, symbol rate and constellation size constrained by average transmitted power and BER. To perform this task, there is a need for BER in AWGN expression for each modulation technique that is easily inverted with respect to rate and power. Unfortunately, exact expressions for BER is hard to find for most prevailing MPSK and MQAM modulation techniques. Therefore, it is a common exercise to make acceptable approximations for the BER upper bounds at high SNRs.

Keeping the above in mind and assuming ideal coherent detection, the BER for MQAM modulation with Gray bit mapping in AWGN channel is bounded by [39, 47]

$$P_b \leq 2exp[-1.5\gamma_s/(M-1)] \tag{5.1}$$

A tighter bound which is good within 1 dB for $M \geq 4$ and $0 \leq \gamma_s < 30$ dB is given by

$$P_b \leq 0.2exp[-1.5\gamma_s/(M-1)] \tag{5.2}$$

The above expressions are merely bounds, which are different from the error probabilities given in Table 3.2. Adaptation designs can make use of the above bounds, which are easily invertible, or other noninvertible bounds by seeking some numerical inversion methods.

5.3.2 Variable-Rate Variable-Power MQAM

The adaptation of transmit power $P(\gamma_s)$ relative to γ_s, constrained by the average power $\bar{P}$ and the instantaneous BER denoted by $P_b(\gamma_s) = P_b$, is considered. Given the received SNR denoted by $P(\gamma_s)/\bar{P}$, the P_b of (5.2) based on γ_s is thus given by [39, 47-49, 53]

$$P_b(\gamma_s) \leq 0.2exp[-1.5\gamma_s P(\gamma_s)/\bar{P}(M-1)] \tag{5.3}$$

Adjusting M and $P(\gamma_s)$ to maintain the target P_b and yield the maximum constellation size

$$M(\gamma_s) = 1 + A\gamma_s P(\gamma_s)/\bar{P} \tag{5.4}$$

where the constant value A is bounded by

$$A = -1.5/ln(5P_b) < 1 \tag{5.5}$$

The spectral efficiency can then be maximised as follows

$$\mathbb{E}[log_2 M(\gamma_s)] = \int_0^\infty log_2\left(1 + A\gamma_s \frac{P(\gamma_s)}{\bar{P}}\right) p(\gamma_s) d\gamma_s \tag{5.6}$$

subject to the following power constraint

$$\int_0^\infty P(\gamma_s) p(\gamma_s) d\gamma_s = \bar{P} \tag{5.7}$$

which concludes the following optimal power adaptation policy

$$\frac{AP(\gamma_s)}{\bar{P}} = \begin{cases} 1/\gamma_{sA} - 1/\gamma_s & , \quad \gamma_s \geq \gamma_{sA} \\ 0 & , \quad \gamma_s < \gamma_{sA} \end{cases} \tag{5.8}$$

where $\gamma_{sA} = \gamma_{s0}/A$ is the optimised cutoff fade depth below which the channel is not used, and γ_{s0} is the cutoff value that can be based on a desired outage probability $P_{out} = Pr(\gamma_s < \gamma_{sA})$ or on a desired target BER above cutoff that is determined by the target BER and $p(\gamma_s)$. The expression of (5.8) is famously known as WF formula.

The adaptive rate and its corresponding adaptive spectral efficiency (ASE), or normalised achievable throughput, are thus given by, respectively

$$M(\gamma_s) = \left(\frac{\gamma_s}{\gamma_{sA}}\right) \tag{5.9}$$

$$\frac{R_s}{B} = \int_{\gamma_{sA}}^\infty log_2\left(\frac{\gamma_s}{\gamma_{sA}}\right) p(\gamma_s) d\gamma_s \tag{5.10}$$

Numerical methods can be adopted to find the solution for intractable (5.10). As such, it can be reduced to a two dimensional real integral with integration kernel e^{-z^2} that can be efficiently computed by using Gauss-Hermit quadrature rules [44, 139], where z is a

dummy variable. The particulars of this method are intimated in Appendix (A) of this work.

Approximating the integrals of (5.10) entails exhaustive numerical computations; simplified tracks for the solution are instead handy. A rather simpler form of power adaptation can be worked out by using the total channel inversion, where the transmitter regulates its power to keep a constant received power. This power control policy of channel inversion results in the following ASE expression [39, 47, 53]

$$\frac{R_s}{B} = log_2\left(1 + \frac{-1.5}{ln(5P_b)\mathbb{E}[1/\gamma_s]}\right) \tag{5.11}$$

where $\frac{P(\gamma_s)}{\bar{P}} = \frac{1}{\gamma_s \mathbb{E}[1/\gamma_s]}$ and implies the tight bound (5.2). Otherwise, the loose bound (5.1) must be used if $M = R_s/B$ is smaller than 4 and hence the term $ln(5P_b)$ is replaced by $ln(0.5P_b)$ in (5.11). The truncated channel inversion can only be applied when the channel is useful and experiencing $\gamma_s > \gamma_{s0}$. The ASE is thereby optimised as follows

$$\frac{R_s}{B} = max_{\gamma_{s0}} log_2\left(1 + \frac{-1.5}{ln(5P_b)\mathbb{E}[1/\gamma_s]}\right) p(\gamma_s > \gamma_{s0}) \tag{5.12}$$

It is worth to mention that the above variable-rate and variable-power (VRVP-MQAM) adaptation technique can be straightforwardly generalised for any other M-ary modulation alphabets. The principle is basically the same; the transmitter power and constellation size are adapted to maintain a given instantaneous BER for each symbol while maximising the data rate. Further details in this regard can be found in [39] and can be summarised as in what follows for the purpose of presentation completeness.

It is required to find the optimal power $P(\gamma_s)$ and rate $R_s(\gamma_s) = log_2 M(\gamma_s)$ adaptation for general M-ary modulation that maximises the average data rate $\mathbb{E}[R_s(\gamma_s)]$ with average power $\bar{P}$ while meeting a given BER target. This can be simplified if the probability of bit error for the modulation is approximated as below [39]

$$P_b(\gamma_s) \approx c_1 exp\left(\frac{-c_2\gamma_s}{2^{c_3 R_s(\gamma_s)} - c_4}\frac{P(\gamma_s)}{\bar{P}}\right) \tag{5.13}$$

where c_1, c_2, c_3 are positive fixed constants and c_4 is a real constant. The power adaptation formula can thus be simplified further as below

$$\frac{P(\gamma_s)}{\bar{P}} = \begin{cases} \zeta - {}^{1}/_{\gamma_s A} & , \quad P(\gamma_s) \geq 0, \qquad R_s(\gamma_s) \geq 0 \\ 0 & , \qquad\qquad\qquad\qquad else \end{cases} \tag{5.14}$$

where the factor ζ is found by solving the average power constraint given by (5.7). While the values of other constants are $c_1 = 0.2$ or 2, $c_2 = 1.5$, $c_3 = 1$, and $c_4 = 1$ for MQAM and the probability approximation will converge to the expressions given in Table (3.2) earlier.

The effectiveness of the above adaptation policy is exemplified in what follows. First is to demonstrate the WF power adaptation principle, which is widely embraced in practical wireless communication systems and can be applied either in time or frequency domains [39, 107]. An illustrative example of using arbitrary transmit and return power levels through few feedback blocks is shown in Fig. 5.2 below. The targeted transmit power has been set to a unitary value, while the returned power levels can be seen fluctuating on levels mostly less than that. Therefore, the transmitter keeps pouring water (power) until levelling up to reach the designated unitary value again, and that is in order to fulfil the requirement of a predefined BER value and below which the performance deems unacceptable. The trend of pumping water (power) also shows a seldom dwelling to zero, which is an indication that the return powers naturally always reside below the required level due to channel degradation effects. After all, this is natural water conduct as it tries to fill all gaps and stabilise surface levels, but the analogy is drawn for power here.

The second illustration is to infer on the performance of achievable spectral efficiency for conventional discrete constellations. The spectral efficiencies so far considered generally insinuate no restrictions on the constellation size; actually, the size is not even restricted to integer values and hence can be continuous. Despite this might look appealing, however, transmission at non-integer rates, though it is possible, the complexity of which is considerably high [39, 40, 47]. Moreover, it is practically challenging to constantly adapt the transmit power and constellation size to cope with the fading channels, particularly of fast-fading nature. Thus, restricting the constellation size to just a handful of values seems to be more reasonable and compelling. Such track is quite safe and does not constitute awkward burdens in real environments.

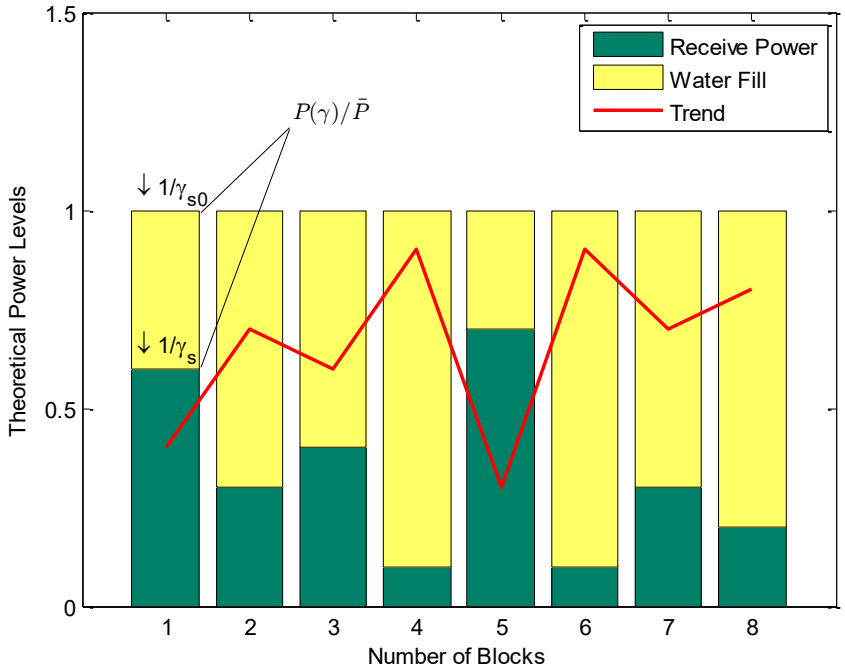


Fig. 5.2 Generic WF operation.

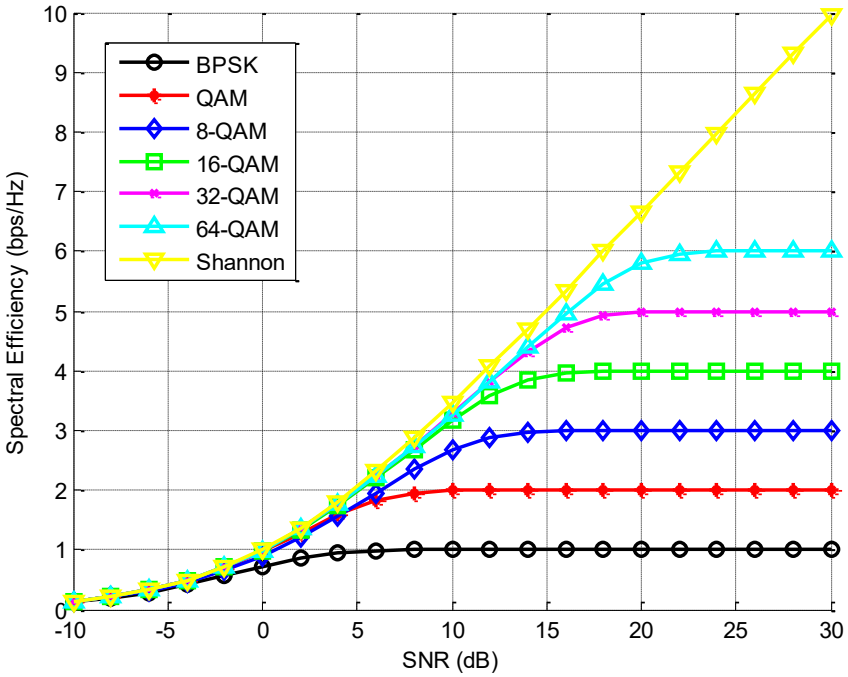


Fig. 5.3 Achievable spectral efficiencies for MQAM signalling.

The achievable spectral efficiencies resulting from applying (5.10) or (5.11) on particular MQAM family members are depicted in Fig. 5.3. It is evident that the spectral efficiency trends saturate because information cannot be sent at a rate higher than $log_2 M$. This is a faithful replica of the theoretical spectral bit rate, or bandwidth efficiency, as illustrated in (3.10) and Table (3.2) earlier. It also can be seen in the low

SNR regime that the reduction in the spectral efficiency and differences between alphabets are negligible. As the SNR increases, the spectral efficiency of equiprobable MQAM constellations tends to asymptotically approach a straight line similar to the spectral efficiency of the AWGN channel. It is worth to note that the approximated integration of the above spectral efficiency terms is accountable to yield the transition probability between closely immediate constellations or alphabets. But the adaptation flavour is still not sensible unless a suitable switching protocol is adopted to allow the transmitter to decide upon particular SNR regions to make inner jumps between desirable alphabets. To accommodate such discrete rate adaptation, particular switching optimisation rules are to be invoked, as will be addressed in the next section.

5.3.3 Constellation Mode Switching

Discrete rate adaptation by restricting the adaptive MQAM to a limited set of constellations is considered in this section. Following the track of [39, 40, 47, 49], some assumptions need to be made in the beginning. A symbol from a constellation in the fixed set $\{C_i^M: i = 0,1, \dots, M\}$, for some arbitrary M, is transmitted at each symbol time. The choice of constellation is determined by the fade level γ_s over that symbol time. The first constellation represents no data transmission as the condition of outage probability P_{out} takes place. For each value of γ_s, the favourable constellation and its associated power level need to be decided. The power varies continually, while the constellation is discrete, and hence this is called discrete-rate variable-power (DRVP).

Assuming FSMC channel, the fade levels can be partitioned into as many SNR segments as required. The constellation size associated to each discretised range of γ_s values can hence be determined. Specifically, the range of γ_s values is divided into M fading segments, $S_i = [\gamma_{s(i-1)}, \gamma_{s(i)})$, $i = 0,1, \dots, M$, where $\gamma_{s(-1)} = 0$ and $\gamma_{s(M-1)} = \infty$. The constellation size M_i is transmitted when γ_s falls in the segment S_i and the spectral efficiency of which is equal to $log_2 M_i$. It is worth to note that the partitioning of the effective SNR depends on the desired BER level for certain QoS. For a reliable transmission to achieve a specific BER target using M_i-QAM, the segment borders are set to the SNR required to meet this desired performance. As the rate is constant in each segment, the target instantaneous BER is achieved by varying the transmit power alone. Such constellation mode selection is shown in Fig. 5.4 for any arbitrary BER target.

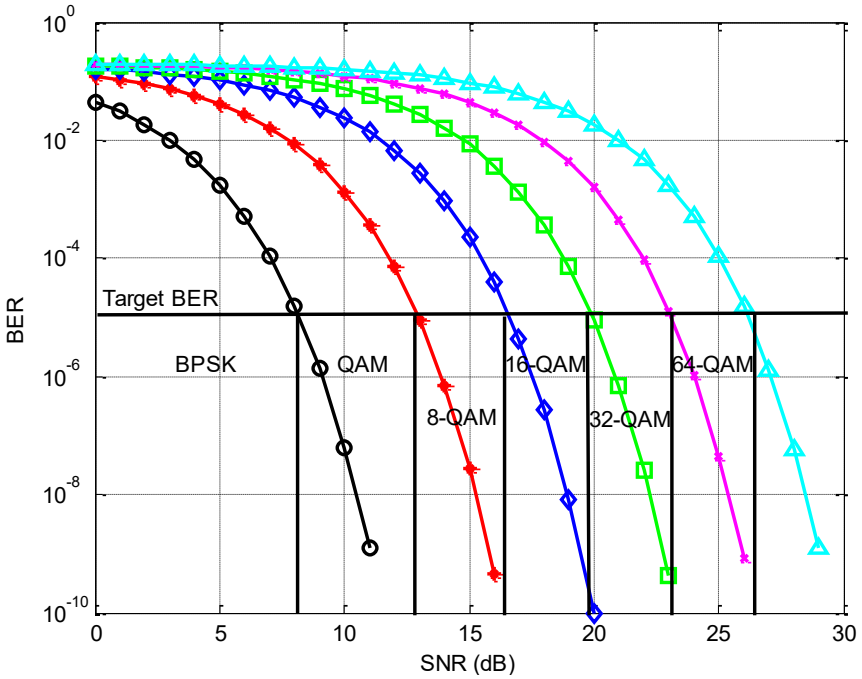


Fig. 5.4 Mode selection according to a target BER.

The design of adaptive MQAM entails the boundaries of the S_i segments be defined. A common practice in determining such boundaries is to probe into possible optimisations suitable to maximise the SE utilisation. However, getting closed forms of such optimal boundaries is not an easy task as it involves exhaustive search methods. Therefore, suboptimal boundaries are generally cheaper and acceptable as they produce the same performance as the optimal ones [39, 40, 47]. Some optimisation methods based on Lagrangian functions to maximise the spectral efficiencies have been explored in [40, 43, 49, 51]. Whereas the LS techniques of curve fitting and utilizing simulated BER for different constellation sizes was embraced for the same purpose in [50].

Recall from (5.9), γ_{sA} is the cutoff fade parameter that needs to be optimised for maximum spectral efficiency value. Therefore, the appropriate choice of γ_{sA} determines the optimal constellation size for each γ_s subject to having no restrictions imposed on the constellations. The DRVP-MQAM scheme performs as follows. For a target BER, obtain the associated γ_s values from the simulated BER curves depicted in Fig. 5.4, or differently in a tabulated form, for each permissible constellation size. Make γ_s partitions with respect to each constellation size and define their boundaries. Compute $M(\gamma_s)$ from (5.9) and allocate within an appropriate partition $M_i \leq M(\gamma_s) < M_{i+1}$ and assign M_i to correspond to a range of instantaneous γ_s values within the same partition.

That is for a fixed range of $\gamma_s$within particular partition, the transmitted constellation size is M_i, which is smaller than $M(\gamma_s)$ to maintain a target BER. For example, if the normalised channel fade lies in the range $2 \leq \gamma_s/\gamma_{sA} < 4$, then the BPSK constellation size of 2 is used for transmission in this case. No doubt, the higher the constellation size, the better spectral efficiency is achieved providing that a clear knowledge of the fading conditions is at hand to render partitioning with respect to the designated BER near optimal.

Once the γ_s partitions and associated constellations are fixed, the power adaptation policy to satisfy the BER requirement and power constraint is to be obtained in the next step. By referring to (5.4) and (5.8), this policy to maintain a fixed BER for the constellation $M_i > 0$ can be achieved as follows [39, 40, 47, 49]

$$\frac{P_i(\gamma_s)}{\bar{P}} = \begin{cases} (M_i - 1)(1/A\gamma_s) & , \quad M_i \leq \gamma_s/\gamma_{sA} < M_{i+1} \\ 0 & , \quad M_i = 0 \end{cases} \tag{5.14}$$

where $\gamma_s \in S_i$, and this power adaptation policy leads to a fixed received SNR for the constellation of size M_i as can be described below

$$\frac{\gamma_s P_i(\gamma_s)}{\bar{P}} = \frac{M_i - 1}{A} \tag{5.15}$$

The desired target BER can eventually be achieved by defining A, MQMA constellation size M_i and returned SNR from the receiver as governed by (5.15). All these adaptive design parameters are commonly arranged in a mapping look-up table [39, 40, 47], which can easily be devised for seven fading segments as shown in Table 5.1.

Table 5.1 Discrete rate and power adaptation for 7 segments.

Segment S_i	**γ_s/γ_{sA} range**	**Constellation M_i**	**$\frac{P_i(\gamma_s)}{\bar{P}}$**
0	$0 \leq \gamma_s/\gamma_{sA} < 2$	0	0
1	$2 \leq \gamma_s/\gamma_{sA} < 4$	2	$1/A\gamma_s$
2	$4 \leq \gamma_s/\gamma_{sA} < 8$	4	$3/A\gamma_s$
3	$8 \leq \gamma_s/\gamma_{sA} < 16$	8	$7/A\gamma_s$
4	$16 \leq \gamma_s/\gamma_{sA} < 32$	16	$15/A\gamma_s$
5	$32 \leq \gamma_s/\gamma_{sA} < 64$	32	$31/A\gamma_s$
6	$64 \leq \gamma_s/\gamma_{sA} < \infty$	64	$63/A\gamma_s$

The overall spectral efficiency for this DRVP policy represents the sum of the individual spectral efficiencies associated with each partitioning segment multiplied by the probability that is γ_s allocated in that segment [39, 40, 47, 49, 50, 53]

$$\frac{R_s}{B} = \sum_{i=0}^{M} log_2(M_i) p\left(M_i \leq \frac{\gamma_s}{\gamma_{sA}} < M_{i+1}\right) \tag{5.16}$$

providing M_i is a function of γ_{sA}, (5.16) can hence be maximised relative to γ_{sA}, subject to the following power constraint, in the same manner of (5.7) but now in a discretised form

$$\sum_{i=1}^{M} \int_{M_i\gamma_{sA}}^{M_{i+1}\gamma_{sA}} \frac{P_i(\gamma_s)}{\bar{P}} p(\gamma_s) d\gamma_s = 1 \tag{5.17}$$

The number of constellations is also restricted for the adaptive total and truncated channel inversion policies and must be admissible from a fixed set of possible alphabets $\{C_i^M : i = 0,1, \dots, M\}$. The spectral efficiencies for these two policies are given below, respectively [39, 47]

$$\frac{R_s}{B} = log_2 \left\lfloor \left(1 + \frac{-1.5}{ln(5P_b)\mathbb{E}[1/\gamma_s]}\right) \right\rfloor_C \tag{5.18}$$

$$\frac{R_s}{B} = max_{\gamma_s} log_2 \left\lfloor \left(1 + \frac{-1.5}{ln(5P_b)\mathbb{E}[1/\gamma_s]}\right) \right\rfloor_C p(\gamma_s > \gamma_o) \tag{5.19}$$

where $\lfloor z \rfloor_C$ means the largest number in the set C less than or equal z.

Tantamount adaptation policies using similar partitioning schemes can also be found in the realm of CR applications [56-58], but those were engaged with channel coding techniques as a vital ingredient of ACM strategies, which will be explored next. Simulation results to corroborate all of the above analyses and the subsequent proposed scheme are conveniently postponed to the end of this chapter.

5.4 Channel Coding Techniques

In the wide framework of wireless communication systems, when the transmission suffers deep fading effects, the channel coding techniques are viewed as a compelling

salvation against such detrimental situations [39, 40, 42, 107, 169], which can be assumed as the third DoF in this work. Remarkably, for the same settings under AWGN, this extra DoF as a reminiscent of coded ACM was claimed to be more power resilient than uncoded ACM in fading channels [49].

By overlaying trellis or lattice codes onto the adaptive modulation schemes, additional coding gain can be attractively produced and so the full ACM concept retained. The subset partitioning inherent to the CM trellis or lattice codes and specifically designed for AWGN channels can be overlaid directly onto the adaptive modulation with the same coding gain approximations [39, 40, 42, 47, 49, 107]. Such feasibility is attributed to the elegant separation ability of coding and constellation design that lies at the heart of CM constructions. The CM is thus considered as a natural choice for the VRVP-MQAM due to the essential independency between the channel coding gain and the modulation. The transmitted constellation power and rate in such an approach can be adjusted relative to the instantaneous SNR without affecting the channel coding gain.

Before elaborating on specific details, let the general ACM principle featuring any kind of channel coding techniques be explained first in what follows. The generic building blocks of ACM complete concept is depicted in Fig. 5.5 [39, 49].

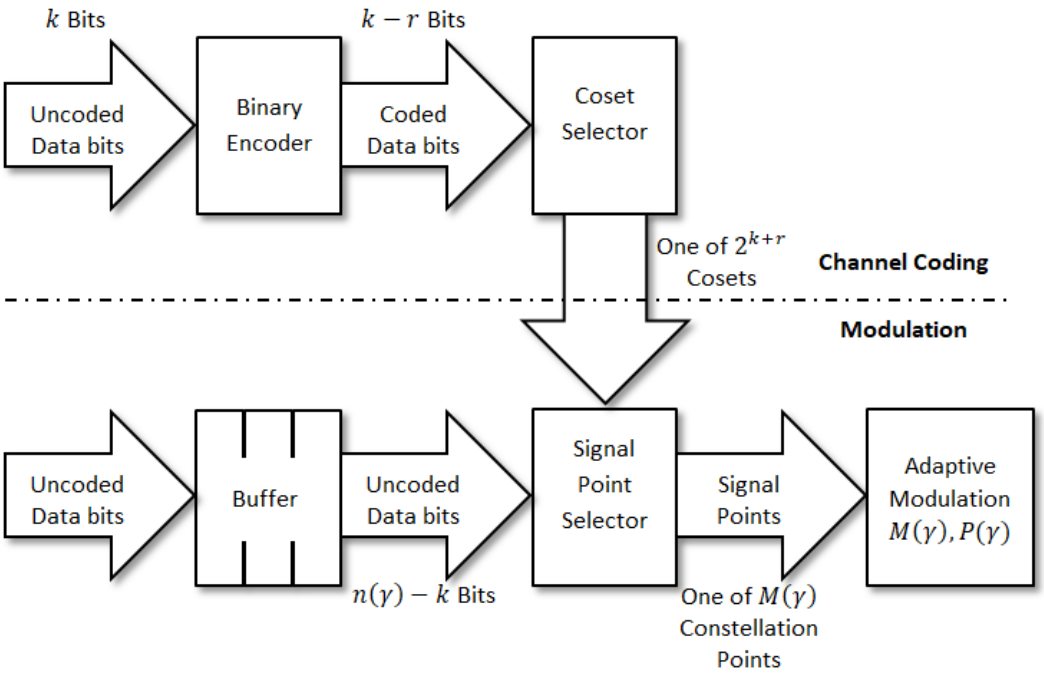


Fig. 5.5 ACM generic diagram.

Let the ACM coding gain denoted by G_c, and for a particular SNR = γ the tight bound of BER expression (5.2) can hence be easily adjusted to be [39]

$$P_b \leq 0.2exp[-1.5G_c\gamma_s/(M-1)] \tag{5.20}$$

and adjusting the constellation size M and the signal power relative to instantaneous SNR and fixed BER target reforms (5.4) as below

$$M(\gamma_s) = 1 + \frac{1.5\psi G_c}{-ln(5P_b)}\frac{P(\gamma_s)}{\bar{P}} \tag{5.21}$$

The number of uncoded bits required to make the coset point selection is given by

$$n(\gamma_s) - 2k/N = log_2 M(\gamma_s) - 2(k+r)/N \tag{5.22}$$

where N is constant and stands for the plausible maximum trellis size, while $n(\gamma_s)$ values vary with time and hence the uncoded bits must be appropriately buffered and synchronised until fitted, as shown in Fig. 5.5.

The optimal power adaptation policy in this case is given by

$$\frac{P(\gamma_s)}{\bar{P}} = \begin{cases} 1/\gamma_{s0} - 1/\gamma_s A_c & , \quad \gamma_s \geq \gamma_{s0}/A_c \\ 0 & , \quad \gamma_s < \gamma_{s0}/A_c \end{cases} \tag{5.23}$$

where the cutoff fade depth is $\gamma_{sA_c} = \gamma_{s0}/A_c$ for $A_c = AG_c$. The CM hence rises the effective transmit power by G_c relative to the uncoded VRVP-MQAM regime. The adaptive rate and the resulting SE can be obtained as follows, respectively

$$M(\gamma_s) = \left(\frac{\gamma_s}{\gamma_s A_c}\right) \tag{5.24}$$

$$\frac{R}{B} = \int_{\gamma_s A_c}^{\infty} log_2\left(\frac{\gamma_s}{\gamma_s A_c}\right) p(\gamma_s) d\gamma_s \tag{5.25}$$

In the next section, the calculation of coding gain is illustrated.

5.4.1 Trellis Coded Modulation

In the early works on information theory, the focus was on reducing power requirements at low spectral efficiencies. This was partially attributed to the fact that no

practical coding schemes existed that could provide meaningful power gains at higher spectral efficiency levels [39, 40, 107]. This changed dramatically with Ungerboeck's discovery of TCM [170, 171] in the early 1980s. In his work, Ungerboeck constructed trellis codes for amplitude modulation (AM), PSK, and QAM schemes. His technique, based on "mapping by set partitioning", as opposed to the conventional Gray mapping used in uncoded modulation, turned out to be the key feature in achieving greater performance gains with TCM compared to uncoded modulation at high spectral efficiencies. One of the important goals of mapping by set partitioning is to insure that parallel transitions in the trellis are mapped into signals far apart in signal space, thus minimising the probability of short, one-branch and error events due to uncoded bits. Such joint optimisation between coding and modulation results in significant coding gains without bandwidth expansion.

Referring to Fig. 5.5, the five elements required to generate the TCM are [39, 107]:

- A binary encoder, whether of block or convolutional type, which acts on k uncoded data bits to generate $k + r$ coded bits.
- A subset (coset) selector that uses the coded bits to select one of 2^{k+r} subsets from a partition of the M-dimensional signal constellation.
- A point selector that uses $n - k$ additional uncoded bits to select one of the 2^{n-k} signal points in the selected subset.
- A constellation map that maps the selected points in N-dimensional space to a sequence of $N/2$ points in two-dimensional space.
- An MQAM modulator; or any other M-ary modulator as applicable.

The fundamental concept behind using the signal-set partition concept in TCM led to noticeable gains due to provision of coding redundancy. Such redundancy is achieved by designing joint coding and signal mapping functions so as to maximise the "free distance" (minimum Euclidean distance (MED)) between coded signal sequences. This allowed the construction of modulation codes whose free distance significantly exceeds the minimum distance of the uncoded modulation signals, at the same information rate, bandwidth, and signal power. Set partitioning has been well adopted as the key for the constructing of efficient CM techniques aimed at bandwidth-

constrained channels. The code distance properties, and hence the coding gain, are determined by the encoder properties and the subset partitioning, which are decoupled from signal shaping.

The common encoder used in TCM is the linear convolutional encoder as the output of which, the binary convolutional codes, is significantly adequate for setting and partitioning [39, 107, 170, 171]. The convolutional encoder of rate $R_c = n/(n+1)$, which is also called a trellis or lattice encoder, introduces forward error correction (FEC) information necessary for recovering the transmitted data. After the convolutional encoder, each symbol has $n+1$ bits, which requires $2n+1$ constellation points. In general, for the same average power, a modulation scheme using $2n+1$ constellation points has higher BER if compared with $2n$ constellation points. The reason for this is that the MED between any two points on a $2n+1$ constellation is smaller, which decreases the noise margin. One prime feature of convolutional encoding lies in imposing a constraint in transforming an n bit input symbol to a $n+1$ bit output symbol. Specifically, it does not allow two consecutive output symbols to be in the eight neighbourhood positions of each other. This results in an increment of the MED between two consecutive output symbols, which is achieved through mapping by the set partitioning method. This process provides an overall performance gain of 3 - 6 dB, depending on the constellation's type and size.

In set partitioning, the M-ary constellation is successively partitioned into $2, 4, 8, \ldots, 2^{log_2 M-1}$ subsets, with size $\frac{M}{2}, \frac{M}{4}, \frac{M}{8}, \ldots, 2$, with progressively larger minimum distances. The assignment of signal subsets to state transitions in the trellis is based on three heuristic rules devised in [107, 170, 171]

- Assign parallel transitions to states of same largest partition. Members of the same largest partition are assigned to parallel transitions.
- Assign adjacent transitions to states of next larger partition, i.e., transitions stemming from, or merging into the same state.
- Use all subsets with equal recurrence in the trellis.

The procedure of set partitioning for 16-state QAM constellation is illustrated in Fig. 5.6 [39, 107, 170, 171], and the description of which follows below.

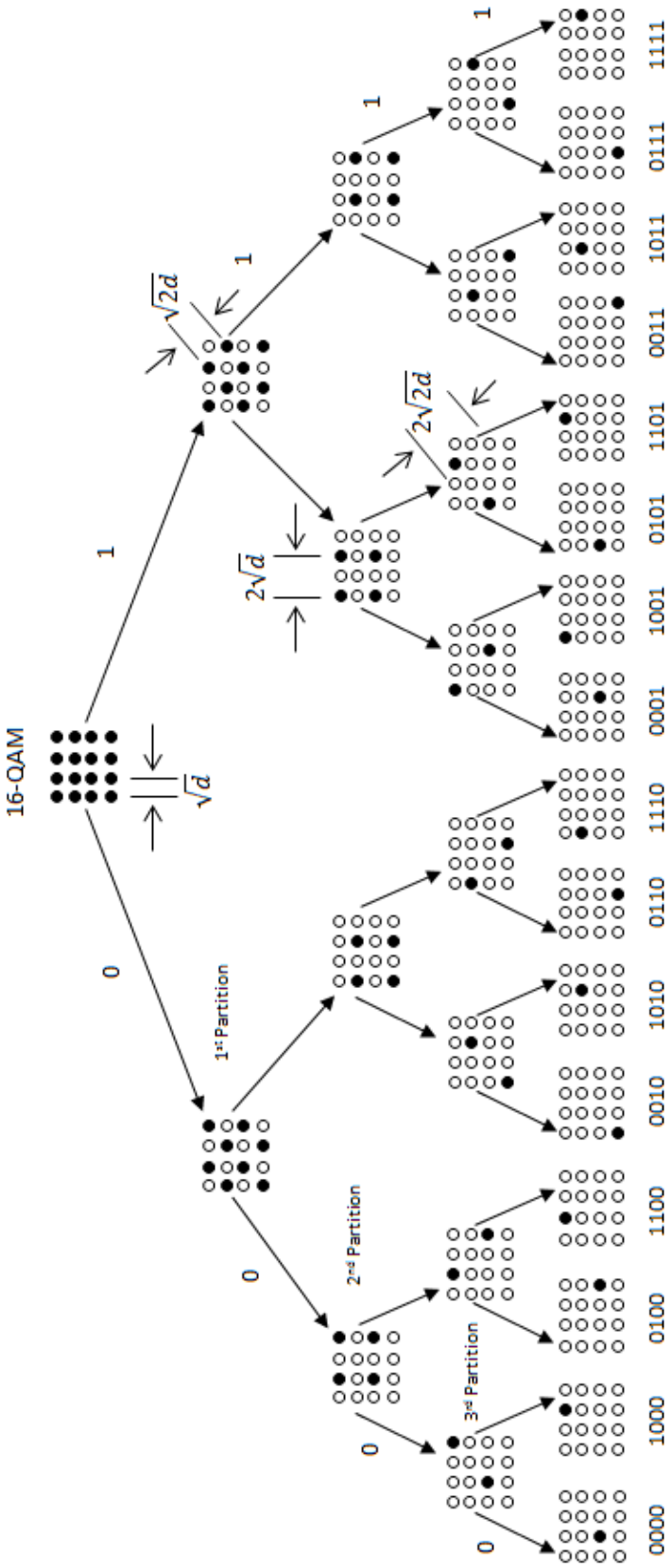


Fig. 5.6 Set partitioning of 16-QAM signal.

If N input bits are considered as a trellis state (or point) in N-dimensional space, then this selection is used to secure some minimum distance d_{min} in the sequence space between possible input sequences. In AWGN channels, the error probability is a reverse proportional function of d_{min}^2. Hence, the BER can be decreased by increasing the separation between each state in the sequence space by a fixed amount. This is called stretching the space, which can be achieved by set partitioning. The d_{min}^2 between points in the subsets is increased by at least a factor of 2 with each partitioning. In the first partitioning, the 16-points constellation is subdivided into two 8-point subsets. The value d_{min}^2 increases to $2d^2$ from d^2. In the second partitioning, each of the two 8-point subsets is subdivided into two subsets of 4-point, and the value distance d_{min}^2 is increased to $4d^2$. This process continues on the subsets until each subset has only two points, and the value d_{min}^2 is now increased to $8d^2$. In QAM constellation, each level of partitioning increases the MED by $\sqrt{2}$. It is the code characteristics that determines the signal partitioning and hence proper codes need to be selected for best BER performance in the TCM scheme.

As indicated earlier, convolutional codes are widely used in many designs of wireless communication systems. There exist different methods for decoding of convolutional and trellis codes and all are based on two known principles; soft decision or hard decision. While ML or MAP algorithms can be used as optimal decoding of convolutional codes, the celebrated Viterbi algorithm (VA) decoding, on the other hand, is trusted for TCM applications [39, 40, 42, 50, 107, 170, 171]. The soft decision of combining VA decoding and ML algorithm is usually tasked to finding the shortest path through one such trellis, where the length of each trellis branch is measured by the distance between a candidate symbol and the signal received. VA decoding is predominantly used for short constraint lengths, while sequential decoding is used for long-constraint codes, where the complexity of VA decoding becomes prohibitive. The choice of constraint length is governed by the desired coding length.

Bearing the above in mind, the effective coding gain can be calculated by [39, 171]

$$G_c = 2^{-2r/N} d_{min}^2 \tag{5.26}$$

where $2^{2r/N}$ is the constellation expansion factor from the r extra bits introduced by the binary channel encoder. Suppose d_p denotes the minimum distance between points

within a coset and d_s the minimum distance between coset sequence, then the minimum distance code can be given by $d_{min} = min(d_p, d_s)$. Table 5.2 illustrates some of the coding gains, or energy savings, that can be obtained by expanding the trellis size in TCM [42, 172]. These are considered for coded 8-PSK (relative to uncoded 4-PSK) and for coded 16-QAM (relative to uncoded 8-QAM). These asymptotic gain approximates can only be achieved at higher SNRs as they decrease when the latter go down.

Table 5.2 Sample asymptotic coding gains of TCM.

Number of trellis states *N*	Coding gain (dB) 8-PSK	Coding gain (dB) 16-QAM
4	3	4.4
8	3.6	5.3
16	4.1	6.1
32	4.6	6.1
64	4.8	6.8
128	5	7.4
256	5.4	7.4

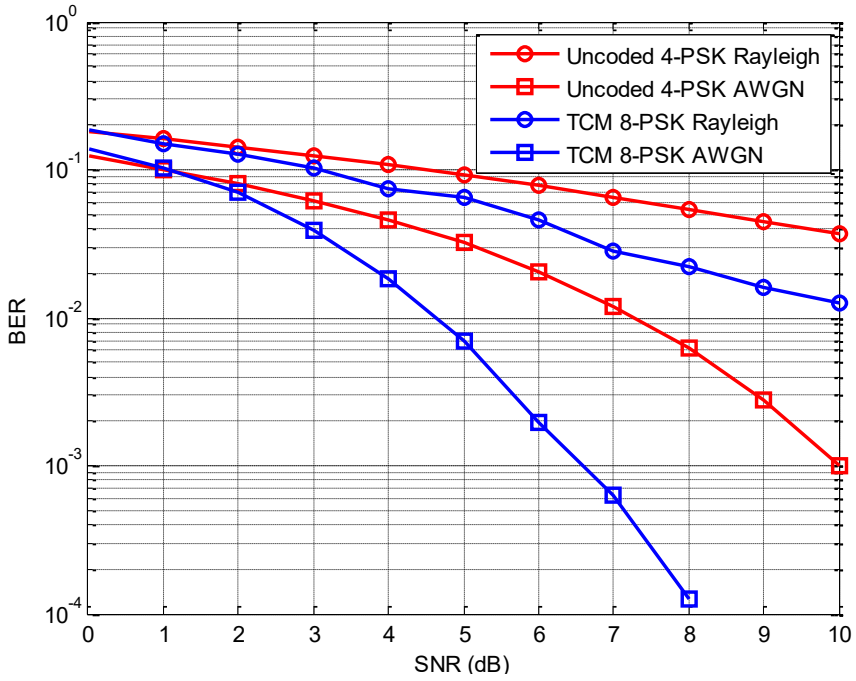


Fig. 5.7 BER performance of 8-PSK TCM and uncoded 4-PSK.

The BER performance of 8-states 8-PSK TCM and uncoded 4-PSK under AWGN and Rayleigh channels relative to SNR increments is shown in Fig. 5.7. Albeit the poor performance is conspicuous for both schemes under Rayleigh channel, however, it is healthier under AWGN channel. Moreover, the BER performance of TCM favourably demonstrates improved trend under both channels and that is primarily attributed to the additional power gains generated by the involvement of sophisticated coding. This exhibits an expressive gain in the range of 3 - 4 dB credited to applying the TCM under both channels, which is in agreement with the range claimed by other studies [47, 49, 170]. However, such gains might not be able to shoot higher relevant to the increase in the number of trellis states as it is constrained by complexity factors [47].

5.4.2 Bit-Interleaved Coded Modulation

Generally, and as can be deduced from the above discussion, the interpretation of adaptation strategies including ACM usually refers to the involvement of TCM schemes. Recap from the above, the trellis structure can be adjusted by fixing the number of information bits entering the encoder, and adapting the number of uncoded bits in accordance to the channel status. If the channel is estimated erroneously, this adaptation scheme is considered not robust to such errors, especially under fading channels.

In order to improve the performance of coding schemes in fading channels, coding design is typically combined with interleaving to offset the effect of error bursts [39]. Powerful designs of binary codes with a good tradeoff between performance and decoding loads have been developed in the last three decades. Among a few is the pragmatic approach of BICM, which is aimed at combining the best out of both aspects; featuring signal-space coding, whilst allowing sophisticated binary codes that virtually can work with any type of modulation formats [39, 40, 107]. BICM avoids the burden of intricate and inflexible design typical of CM, and has eventually established itself as a de-facto standard for high bandwidth efficiency and power efficiency in a wide range of communication systems.

The strategy of BICM is well perceived as an attractive option because it increases the time diversity, or equally the code diversity [39, 40, 42, 52, 107]. In BICM, the code

is left fixed, while the signal constellation is adapted as per the channel conditions and hence the rubric "adapting the constellation size") was proposed [42]. Previous studies also indicated that the code structure of BICM can be readily extended to adaptive systems that support higher mobility platforms [42, 52]. Moreover, the combination of BICM-OFDM was also alleged to be a good candidate for power line communication (PLC) systems.

The paradigm of BICM was firstly coined to the landmark work of Zehavi [172], and its information-theoretic analysis was then extensively studied by Caire et al. [173], see also a recent study of the latter in [174]. Zehavi realised that the code diversity, and hence the CM reliability over fading channels, could be enhanced furtherly. His idea was embarked on making the code diversity distinctly equal to the smallest number of information bits, rather than channel symbols, along any error attribution [45, 172, 173]. To this aim, a bit-level interleaving at the encoder output and an appropriate space diversity bit metric as an input to the VA decoder was employed. This in fact allowed the channel encoder and the modulator to be separated at the bit-level interleaver and hence chosen independently. This is exactly what made BICM so popular in rendering simple and flexible designs for wireless communication systems. BICM has favourably stemmed as a robust structure, since the receiver has little to worry against channel variations. Increasing the Hamming distance, which is equivalent to the code diversity, is championed by BICM and thus lends itself as dominant in fading and AWGN channels as well [174].

To obtain large M-ary Hamming distance for high rates, the complexity required by TCM is much larger than that required by BICM. The Hamming distance in TCM represents the number of coded symbols that differ between different code words or error events, while in BICM it is given by the differences of the underlying binary codes [107, 174]. Hence, for a given complexity, BICM compares favourably with respect to TCM in fading channels, especially for high rates. Despite that, it comes at the price of a marginal reduction of the Euclidean distance, and hence some slippages in the coding gains could be encountered; however, the BICM performance is still considered superior to that of TCM.

An extra bonus of BICM compared to TCM, or other schemes, is the straightforward implementation of ACM systems, which is attributed to the flexibility

imposed by the bit-level interleaver [52]. In order to increase the spectral efficiency, BICM can be combined with high-order modulations such as the common MPSK and MQAM. In the view of iterative processing, further elevation in the performance of BICM is also feasible by allowing the exchange of information between the demapper and the decoder. This scheme called BICM with iterative decoding (BICM-ID) was developed in [175] and explored in many other studies and only few can be mentioned here [176, 177].

The BICM system studied by Caire et al. [173] is depicted in Fig. 5.8 below. Similar replicas can also be found in [107] and with feedback link in [176, 177], while an adaptive BICM modelling was addressed in [52]. The original system ingredients advocated by Zehavi [172], and equally with feedback link in [175], were slightly focused onto the 8-PSK track, yet the generalised foundation of BICM was well perceived in his study.

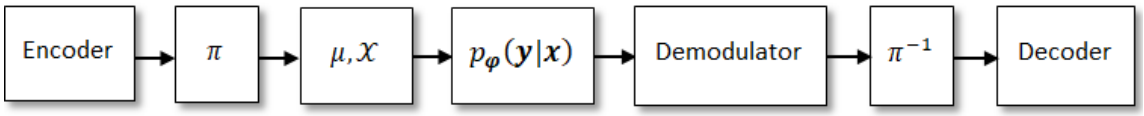


Fig. 5.8 BICM system diagram.

It is worth to know that the system in Fig. 5.8 applies evenly for BICM and any other CM in general [173], but the differences lie in the details of each block and in what follows is a description relevant to BICM. The building blocks of the above BICM scheme consists of 1) an encoder; 2) an interleaver π; 3) modulator, modelled by a labelling map μ and a signal set $\mathcal{X}$, i.e., a finite set of points in the complex N-dimensional Euclidean space $\mathbb{C}^N$; 4) a stationary finite-memory vector channel whose transition probability density function $p_{\boldsymbol{\varphi}}(\boldsymbol{y}|\boldsymbol{x})$, $(\boldsymbol{x}, \boldsymbol{y})\mathbb{C}^N$ may depend on a vector parameter $\boldsymbol{\varphi}$; 5) a demodulator; 6) a branch metric deinterleaver π^{-1}; 7) a decoder.

The design of BICM can be summarised as follows [173]

- Given a signal set $\mathcal{X}$ of M-ary cardinality and dimension N, a desired decoder complexity τ, and a required spectral efficiency, choose the best binary convolutional code of rate $R_c = k/n$ such that spectral efficiency is equal to $R_c log_2 M/N$, with complexity $\leq \tau$.

- Depending on the application (AWGN, Rayleigh fading, etc.) and on the detection technique (perfect CSI, no CSI, etc.), choose the labelling map μ according to some optimality criterion. For perfect CSI, practical and sensible criteria are the minimisation of the average number of signals at distance d_{min} in the complement subset of a signal $\boldsymbol{x}$ and the maximisation of the harmonic mean of squared MED between complementary subsets of $\mathcal{X}$, which is denoted by d_h^2.

Table 5.3 below summarises the bounds on the MED and on the minimum code diversity of BICM and TCM, for 16-QAM codes with $R_c = 3$ bit/dim [107, 173]. As can be expected from the above discussions, BICM achieves better code diversity, while TCM achieves better MED for the same complexity overheads. Further into the BICM performance assessment, the BER results of 8-states 8-PSK BICM and uncoded 4-PSK under AWGN and Rayleigh channels with respect to SNR variations are plotted in Fig. 5.9 below.

Table 5.3 MED and code diversity of BICM and TCM for 16-QAM with $R_c = 3$ bit/dim.

ENC	**BICM**		**TCM**	
Memory	d_{min}^2	Code diversity	d_{min}^2	Code diversity
2	1.2	3	2	1
3	1.6	4	2.4	2
4	1.6	4	2.8	2
5	2.4	6	3.2	2
6	2.4	6	3.6	3
7	3.6	8	3.6	3
8	3.6	8	4	3

By comparing the BER trends of TCM and BICM in Fig. 5.8 and Fig. 5.9, respectively, the superiority of the latter is evident especially under Rayleigh channel conditions. Such reward is a direct product of the enhanced Hamming distances and thus the code diversity orders. It was indicated that such favourable BICM performance is attributed to its ability to increase the time diversity, or code diversity, of systems operating over fading channels by providing an independent fading component for each channel bit out of the convolutional encoder, as opposed to each channel symbol, as in standard symbol-interleaved TCM systems [52]. Furthermore, the BICM structure has

even greater potential as it does not require the use of uncoded bits to adapt constellation size. The uncoded bits in TCM can only achieve a diversity of one, but the inherent diversity of BICM protects every information bit from channel prediction errors. Apart from that advantage, however, the BICM action is slightly inferior to TCM under AWGN channel. These results are in agreement with those obtained earlier in [39, 40, 52, 107, 170-177].

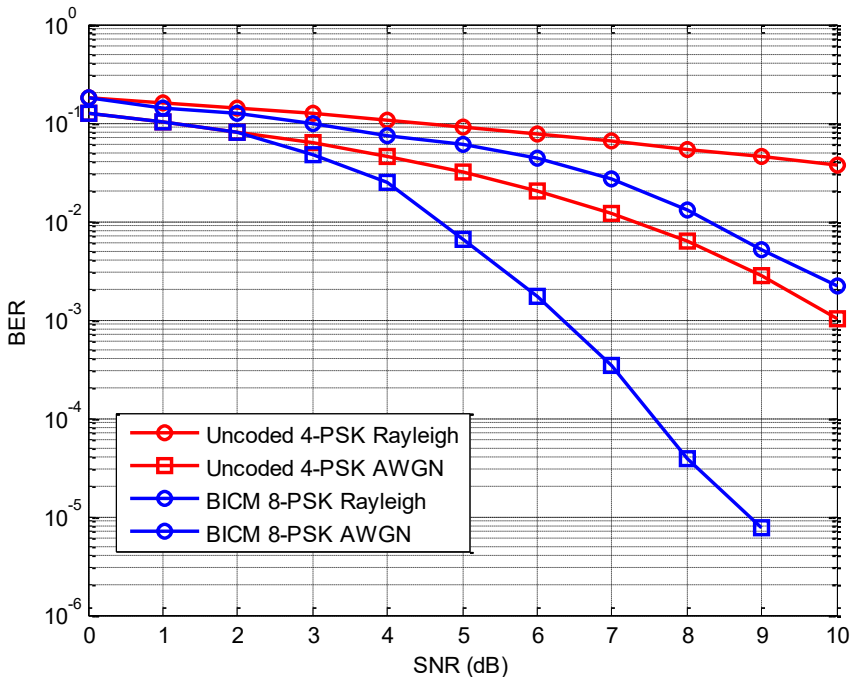


Fig. 5.9 BER performance of 8-PSK BICM and uncoded 4-PSK.

5.5 Proposed Adaptive System

The essence of having a resilient CSI estimate and return it back for the transmitter to adjust its operational parameters has been well justified in the above discussion. It is worth to note that adaptive transmission, based on accurate CSI estimates at the receiver and a reliable feedback path between the receiver and transmitter, was first proposed in the late 1960's [47]. However, interest in these techniques was short lived, possibly due to hardware limitations, lack of good CSI estimation techniques, and/or systems focusing on point-to-point radio terminals without feedback links. The fact that these issues are less constraining in current technologies, coupled with the growing demand for communication systems of high SEs, has revived interest in adaptive modulation

methods. There is no short supply of various strategies of CSI estimates and perceptive feedbacks in the build-up of adaptive systems in the literature. Recall that some techniques have already been explored in Chapter Four of this work, but a quick recap would be useful in such situation.

5.5.1 Technical Overview

The section provides a digest on the technicalities that are meaningful to grasp the incentives behind the proposed adaptation system and engage the most suitable tools and assumptions for analysis thereto. To aid adaptive QAM modulation schemes, the premises of current and previous CSI estimates for channel modelling and to predict future CSI parameters was addressed [43]. The regime of block-fading to enable CSI estimation schemes based on training sequences and then updating the transmitter to perform the necessary adaptation was explored [44]. Various instrumental CSI estimation and equalisation techniques were investigated [45, 46], where involving pilot symbols (or PSAM) was considered valuable when the system is of an adaptive type [46]. A perfect predicted or estimated channel at the transmitter was assumed in the design of adaptive modulation systems [47, 53]. Such an assumption makes the problem formulation and analysis more tractable. An unbiased quadratic regression of the past noisy channel estimates was utilised in the channel predictions [48, 50]. While the adaptive BICM as a robust tool against estimation errors in the current fading channels was motivated in [52].

Moreover, imperfect feedback returns between transmitter and receiver pose tough challenges on the design of ACM systems. The impact of such degraded links, which may constitute delays or error-prone channel estimates, on the performance of ACM systems was addressed in [47, 52, 158, 178]. Analysis has shown that the BER performance in VRVP technique is sensitive to estimations of both CSI and feedback path delay and this needs a special consideration in any practical enactment [47]. It was also noted that when the CSI estimate is outdated, the distribution of the current channel fading conditioned on the estimate has a Rician distribution [52, 178]. Furthermore, it was shown that the effective channel conditioned on estimates of outdated fading more likely reveals Rayleigh attributions as the mobility of terminals increases, and this accordingly greatly impedes the ability to use uncoded bits for rate adaptation [52, 178].

The TCM approach in such cases fails to be an efficient CM strategy because errors from poor predictions dominate the BER performance in any wireless communication system, whether adaptive or not. Consequently, this made the BICM schemes healthier candidates for TCM replacement, chiefly in Rayleigh channels.

In practical reality, the feedback information is commonly quantised to limit the return channel bandwidth [48]. Generally, there are three styles of exploiting partial CSI at the transmitter when the feedback channel is assumed error free with limited capacity, namely; channel vector quantisation, scalar quantisation, and quantised signal adaptation schemes, see [158, 159] and the references therein. The channel quantisation directly applies vector quantisation (VQ) on the measurements at the receiver using the traditional MSE distortion metric, and then performs signal optimisation at the transmitter conditioned on the quantiser output [158]. In most cases, this invokes a very low-capacity requirement on the feedback channel. The receiver is to perform CSI estimates and informs the transmitter through the use of a low capacity feedback link with probably some kind of delay [179]. For the maximum mobile speed of interest is 100 km/hr and the corresponding Doppler frequency $f_D = 200$ Hz at 2 GHz band, the required feedback channel capacity was found to be 7.2 kbps, which is much lower than the forward capacity [180]. Reduced feedback channels using long-rang fading prediction was also suggested to insure performance gains and reliable adaptation in mobile radio systems [181]. An expansion upon previous studies on the performance of cellular systems with different link adaptation strategies revealed that AMC is the most promising in terms of overall performance [182].

Bearing the above in mind, it is well understood that transmission errors are generally unavoidable in any practical system. But the solution to this dilemma is available, which is handled by the retransmission protocol known as ARQ, where the receiver uses a feedback channel to inform the transmitter about a successful decoding - via a positive acknowledgment (ACK) message - or about a decoding failure - via a negative acknowledgment (NACK) message, which triggers a new transmission round (or a retransmission) [44, 54, 183]. Increasing the number of retransmissions improves the chances for successful decoding. The HARQ, on the other hand, is also accessible and in which coding is done across the transmission rounds, and thus is intimately related to the AMC for which coding and modulation are the core elements [54]. The NACK process is inherent to the VA algorithm [183]. It is realised by monitoring the

difference between the survivor path and the path that has the next best metric in the decoding process. It is also not a surprise to reclaim, as was partially explained earlier in the context, that the FSMC channel modelling directly lends itself for ACM systems [183], and the IMM algorithm provides one way of coping with such cardinality.

On the other side, link adaptations based on MIMO-OFDM technologies have been well examined in the various studies, and only a few can be enumerated here for the purpose of exposure consistency [184-187]. It was generally noted that as long as the effective multipath delays in OFDM systems are within guard intervals, orthogonality between subcarriers still can be secured [159]. Eventually, any performance loss due to imperfect CSI disturbs distinct subcarriers only, i.e., wrong decisions do not spread.

A particular link-quality metric for MIMO-OFDM, mean mutual-information bit mapping to be used for fast link adaptation was proposed in [184]. Their results indicated that introducing a correction parameter for this metric considerably improves the accuracy of the pairwise error rate (PER) estimation and they were mainly exemplified for the IEEE 802.11n standard. An adaptive codebook-based CSI prediction and interpolation scheme was proposed for MU MIMO-OFDM systems [185]. In that scheme, geodesic CSI prediction was employed at the receiver to mitigate the feedback delay effect and geodesic CSI interpolation was performed at the transmitter to mitigate the clustering feedback effect. On the other hand, the study in [186] can be well perceived as an extension to the study in [47] as the full BER analysis based on MIMO channel estimation was furtherly elaborated. The main results of [185] revealed that the CSI prediction error degrades BER in SM MIMO systems much more than BF MIMO systems due to the eigenchannel interference that arises from coupling. The results of [187] concurred with those given in [186], and demonstrated the resilience of adaptive PSAM in transmit BF, or MRT, design to both channel estimation and prediction errors.

Having said all that about adaptation requirements in legacy wireless communication systems, it becomes now imperative to probe into their association with CR systems, yet again the number of studies in such vital area is so dense and only a few are attempted here. The study in [57] explained that it was as early as 1968, Hayes verified that one way of efficient mitigations to combat the harmful effects of channel fading is to adaptively adjust the modulation and/or the channel coding format in

addition to other system parameters based on the quality level of near-instantaneous channel perceived by the receiver, which is aided to the transmitter through a feedback return path. In the general context of CR though, a common approach is the individual SU or CRN to exploit a PU link state feedback, monitor how this changes because of the CRN operation and thus estimate the SU-to-PU channel gains. All studies in [56-58] conceptually corroborated that such requirements are very central for the CR systems to achieve high spectral efficiency targets and better exploit the permissible spectrum. The following were considered, eavesdropping the PU feedback channel and detecting the ACK/NACK packet [56], perfect channel estimation [57], and the near-instantaneously adaptive switching mode that entails a reliable feedback link [58].

In the context of CR systems, the sensitive criticality of channel and interference uncertainties has been well accentuated, especially for underlay operation [98]. A channel's estimation based on the received PUs' signals is a widely used approach of orchestrating SUs' transmissions. Most recently, PUs equipped with the AMC capability similar to SUs is a radical jump in many studies [188-191]. The AMC capability of PUs not only improves the primary-link capacity and robustness but also potentially enables SUs for better spectrum inspection and access prospects. That is to be done obliviously by allowing PUs to change their AMC modes while sustaining their average performance [188]. The obliviousness here means that PUs change their AMC mode only according to the signal-to-interference-plus-noise ratio (SINR) that it perceived, without any need for explicit SUs notification. It was also shown that underlay SUs heavily rely on knowing the CSI of interference links and the acquisition of CSI via channel estimation [189]. The cooperation of PUs greatly contributes in the minimisation of overheads in such links' processing, while the SUs' transmission in the overlay mode mainly depend on the PUs' traffic and reliable SS in contrast to underlay. The performance of spectrum sharing using Alamouti OSTBC under correlated antennas combined with several adaptive transmission techniques over Rayleigh fading channels was presented in [190]. The assumption of SUs having the CSI knowledge of PUs' links was attended by employing perfect channel estimation methods. Given the superiority of BF, or ET as typified earlier in the context, in achieving robust transmission adaptation relative to other kinds of MIMO configurations in CR systems, a notable survey can be accessed in this regard [191]. A wide range of closed-loop control characteristics in CRNs has also been illustrated in this survey study.

5.5.2 System Design

Building on the all factual foundations discussed above, the design of proposed ACM system has now come to the fruition stage. The envisioned system tailors the essential components to help be robustly become immune to changing channels and thereby enhances the overall performance in terms of the specified spectral efficiency, BER and QoS levels.

Guided by the various models of ACM systems [57, 179, 180, 181, 183, 186, 187, 191] and after revisiting Fig. 4.3 and Fig. 5. through all their cross-referenced representations, the proposed ACM system layout taking into account the essential components of three DoFs can be devised as depicted in Fig. 5.10 below.

The input data to be transmitted in the system of Fig. 5.10 is generated by a data source and coded by either TCM or BICM, incorporating punctured convolutional encoders of variable rates, to combat the nuisance effects of channel fading. The coded sequence is then superimposed on classical M-ary constellations employing the VRVP adaptation scheme. This will accordingly vary the achievable spectral efficiency level for each individual constellation as per the preassigned BER levels. The overall ASE performance can then be obtained by applying the switching mode policy. This is done by segmenting the estimated SNR values relative to the targeted BER for each single mode constellation; weighing all modes as per their individual probability of occurrence and then adding all to get the final result. The resulting near-instantaneous ASE, orchestrated by the VRVP and switching mode engines, has a strong reliance on the SNR variations inferred at the CSI estimate output.

As also can be seen from Fig. 5.10, the innovative FET for CSI estimate, proposed earlier in Chapter Four, has well been exploited in the new ACM paradigm. Owing to the extensible virtues of BF, or ET, featuring the FET structure has been proven to be a very attractive option. That is directly credited to the pivotal power of FET in creating a great saving in the overall system complexity. Complex MIMO channels readily surrender to the SVD process and decompose into separate parallel channels. The FET is thereby capable of quantifying only one scalar variable, which represents the total sum of the most dominant eigenchannels. Such ultimate desire can be easily achieved

without sacrificing the correctness of analytical approximations that heavily rely on such space reduction.

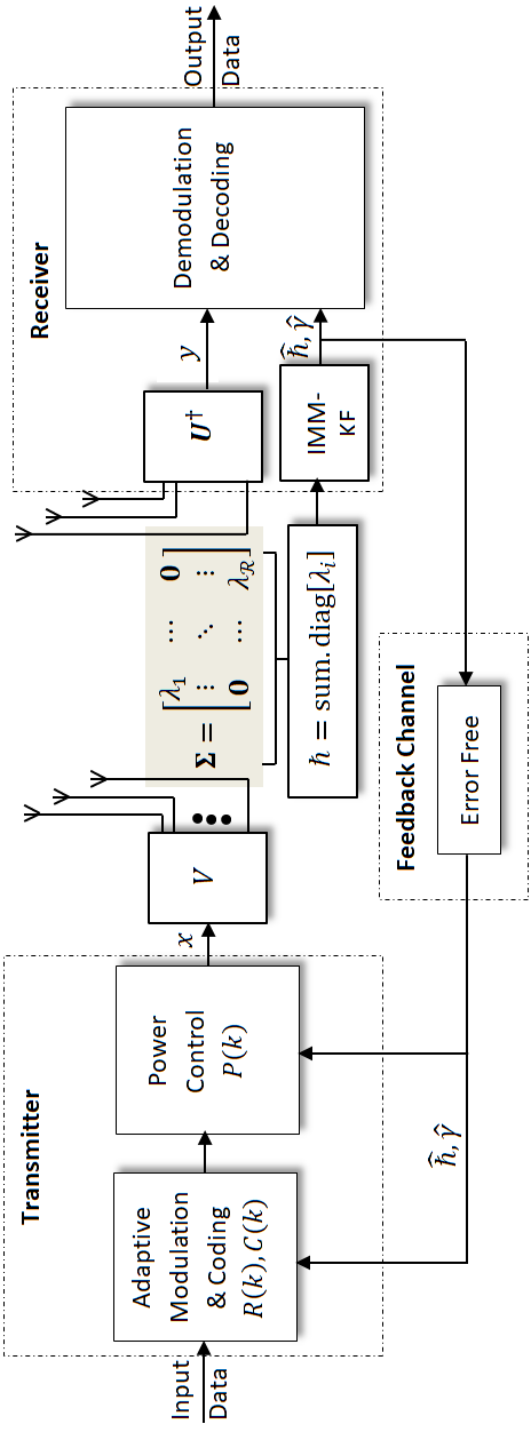


Fig. 5.10 Proposed ACM system model

Driven by the FET idea, the scalar IMM-KF is built to retain reliable CSI estimates against varieties of channel fading and the ACM performance can hence be lifted to a

higher level. That is because the IMM-KF amalgamation is very responsive against slow and rapid channel changes and can span a wide range of dynamical operation. The use of ET and eigenchannels, whether given or estimated, is not new and has been fairly investigated for ACM applications in the literature [44, 158, 159, 186, 187, 191]. But to the best of the author's knowledge and after thoroughly inquiring such a claim, using the IMM-KF to estimate the FET is a unique approach and has never been experimented before. More on that, the approach of summing the eigenmodes was only inflicted in a very few studies such as [159] and [192] and the references therein. In [159], a particular metric for the SINR of each eigenmode is identified and used in the adaptive signalling to ensure a robust adaptive algorithm to account for CSI imperfections and practical operating conditions explicitly. The totality of adaptive throughput was thereby summed over all effective eigenmodes with respect to their probability weights. While in [192], the problem of maximising sum rate of a multiple-antenna Gaussian broadcast channel was addressed. Their approach was based on adding up the eigenmodes of MU transmitting over a MIMO channel and then performing adaptive WF. The IMM-KF based on the holistic FET can virtually be applied to the same above situations, but with great complexity and processing overhead savings.

The calculation of the SNR, expressed in (3.6) and (3.7) of this work, can be done by invoking the estimated FET as will be shortly seen in the next section. Such estimated and calculated parameters have to be returned back for the transmitter to inform about the status of forward channel and other parameters. Such backward update is typically implemented by using an error-free, zero-delay and secured feedback channel. Such an ideal feedback channel and transmitter reception seems fairly reasonable and practically feasible in reality [43, 48, 50, 51, 53, 54, 176, 183, 187, 193]. That is true since the feedback channel is typically quantised and is of lesser bandwidth. Despite such idealism might not be well absorbed in some situations, especially when the return link has errors below certain wall barriers [98], however, these constraints can be easily mitigated by having channel predictions on either transceiver sides [43, 48, 50, 186, 187]. Typically, the prediction span is assumed to be long enough to compensate the delays in the adaptation control loop.

The issue of channel prediction is crucial to adaptive MRT systems, especially those employing PSAM [187]. That is because the receiver of such systems should not only

estimate the MIMO channel based on PSAM for coherent detection and MRC purposes, but also predict the channel ahead of time, as a counter-measure to the feedback delay. That accordingly renders the transmitter to perform two adaptation tasks, namely; selecting the optimal BF vector and then performing the VRVP while adhering to a pre-specified BER value [187]. Referring to [193] and its references, it is shown that the BICM-ID using soft feedback converges to the EFF performance, and the degradation of hard-decision feedback over soft-decision feedback is small when mixed labelling is used. Albeit all above debates, this work adopts a perfect return channel for ease of exposure.

5.5.3 SNR Estimation

The topic of SNR estimation is of a paramount importance and has been extensively researched in countless communication systems. Among many applications, the SNR estimation is typically employed in power control, mobile assisted handoff, and adaptive modulation schemes, as well as soft decoding procedures, to name just a few [194-196]. The rapid development of these applications in the last two decades has triggered an intensive pursuit for SNR estimators that possess better accuracy and less complexity. The main interest in the SNR estimation was actually started in the mid-1960s, when Kerr and Gagliardi handled this issue in relation to the transmission of pulse-code modulation (PCM) through coherent channels [197, 198]. It turned out that the benchmark of most SNR estimators is the MLE, which reduces to the optimum MVUE when the SNR $\geq$ 4 dB [194-198]. Such estimators are naturally heuristic and can be easily defined and calculated.

The SNR estimation not only involves the legacy wireless communication systems, but also has recently induced great potentials in the design and development of new CR systems [98, 74, 76, 78, 199]. These sample studies featured the noise power estimate using the MLE algorithm, or the MoM on a number of occasions, for the AMC and SS applications in CR systems. Such estimated noise power can then be utilised to calculate the SNR as per the expressions outlined earlier in the context. By carefully examining the algorithms in the above given studies [74, 76, 78], it can be easily deduced that they are equivalent to the replicas originally developed by Kerr and then elaborated by Gagliardi [197, 198]. The results in [199] are an exception, as the

derivation of the noise power estimator was formed using the renowned Yule-Walker method.

From another perspective, an overview of the state of the art on channel quality metrics focusing mainly on interweaved CR systems was addressed [200]. The investigated class of channel quality indication metrics comprised the CSI estimate quality and without regard of fading rates and size of the signal constellation. In addition to delivering reliable indicators under noise and interference, these metrics encompassed BER, SINR, outage probability, and channel quality estimation index. The main conclusion derived from that study is that the SNR parameter is so crucial that it impacted on each and every performance metric. This is why the issue of SNR needs special attention while having a direct exposure or indirectly devoted through pertinent aspects.

Galvanised by the FET term, the noise power can be explicitly estimated by subtracting the estimated signal power from the total received power as follows [201]

$$\hat{\sigma}_n^2 = \frac{1}{K}\sum_{k=1}^{K}\left(|y(k)|^2 - \left|\hat{h}(k)\right|^2\right) \tag{5.27}$$

and the SNR estimate can then be computed by applying $\hat{\gamma}_s = \hat{\sigma}_{\hat{h}}^2/\hat{\sigma}_n^2$. That is assuming the signal power is known and of unity value. Otherwise, the signal needs to be estimated first, as will be seen in the upcoming chapter on AMC, and then along the channel and noise estimates to compute the SNR value. The expression (5.27) fairly assumes that the noise is of WSS type as well, and hence its behaviour does not significantly change over K samples while negotiating both estimates. Once the value of $\hat{\gamma}_s$ is found, it should be engaged in all of the ACM expressions described earlier in the context, and as can be similarly inferred from the analysis in [43, 48, 50, 159, 186, 187], whether fully fledged or embodied by the most influential eignemodes.

5.6 Simulation Results and Discussion

This section is devoted to assess the performance of the proposed ACM system based on the IMM-KF structure and by using simulation methods. The assumptions

made are; the forward channel is slowly varying with unity power and the backward channel is ideal.

In the first exercise, suppose an active PU is sending a BPSK sequence of 129 samples in the direction of a Rayleigh fading channel. A SISO scheme with effective SNR states of -10 dB, 0 dB, and 10 dB values to embody the channel variations are used. These are the FSMC states with assumed local, neighbour and remote transition probabilities take on the values of 0.8, 0.2 and zero, respectively. These transition probabilities are also initially assigned for 3-bank KFs to facilitate the IMM algorithm. Each KF has a single σ_n^2 value, such as 0.01, 0.1 and 1, to correspond to mild, moderate and harsh channel conditions, respectively. The signal and AWGN values and their power ratio are assumed fixed and normalised. The values $\hat{h}_{0|0} > 0$ and $P_{1|0} \geq 1$ can also be arbitrarily initialised.

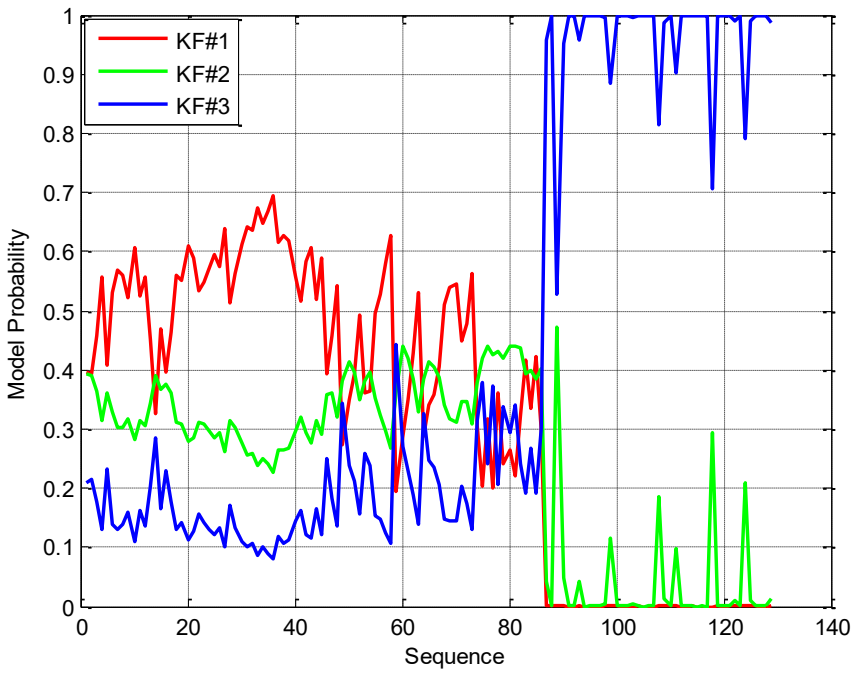


Fig. 5.11 Mode probability of 3-bank KFs.

The receiver on a SU side, knowing the PU signal and armed with the IMM-KF as per the above settings performs the CSI estimation task. The IMM-KF performance in terms of mode probabilities and CSI estimate magnitudes is reflected in Figs. 5.11 and 5.12, respectively. As the strength of the observed channel increases, the IMM

automatically switches between KFs of different internal parameters to cope with such variations. While it is important to realise that the total summation of all mode probabilities is equal to one, the trend shown in Fig. 5.11 clearly indicates when the switching event is taking place. Fig. 5.12 depicts the observed and estimated channel strength, which is in conformance with the initialised SNR range mentioned above. Fig. 5.12 also demonstrates the robustness of IMM-KF in achieving an accurate estimate. This estimate is calculated using (5.27) and after inserting in $\hat{\sigma}_h^2/\hat{\sigma}_n^2$ yields the SNR value $\hat{\gamma}_s$. Therefore, Fig. 5.12 can be interpreted as a faithful manifestation of the SNR behaviour as well.

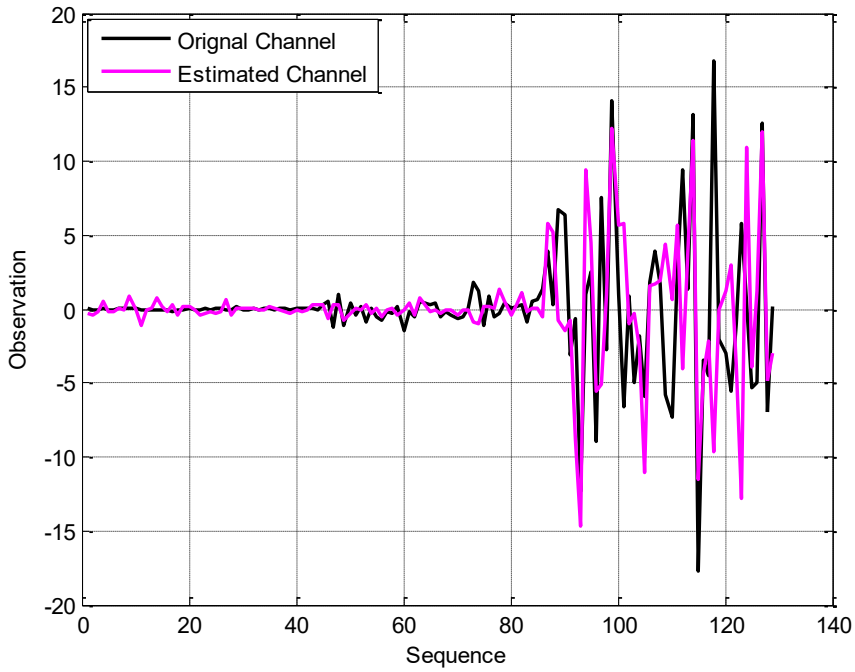


Fig. 5.12 Observed channel and its estimate.

The second exercise is meant for the WF adaptation in various MIMO channels with Rayleigh fading effect. The same above given simulation settings are used, but now the transmission is taking place over roughly 40 blocks, or packets, and each composes of about 20 samples. This is to resemble the effect of a Rayleigh block-fading channel that is slowly varying from one block to another. The constraint of this exercise is denoted by the average power $\bar{P}$ of unity value and the instantaneous BER of 10^{-3}. The functional results of this adaptive WF scheme for different MIMO combinations are depicted in Fig. 5.13 above. As expected, the results match the fictitious one

illustrated earlier in Fig. 5.2, but the results here are taken from simulation setting that could be encountered in reality. Fig. 5.13 (a) for 1×1 illustrates how much power is needed to pour into the transmitter in order to equalise the channel loss as it is being interrogated and reported back to the transmitter to make appropriate adjustments. In as much as the number of MIMO antennas increases, the better performance is discerned and hence the less power is poured as can be seen in Fig. 5.13 (b), (c) and (d). That is certainly attributed to the benefit of spatial diversity gained by having many antennas on both the transmission and the reception sides. Generally, the power adaptation task has been well taken care of by the transmitter after getting the necessary report on the CSI estimate from the receiver, and both collaborate to achieve the predefined requirements. It is also worth to mention that the proposed IMM-KF based on the FET approach is performing appropriately in such scenarios.

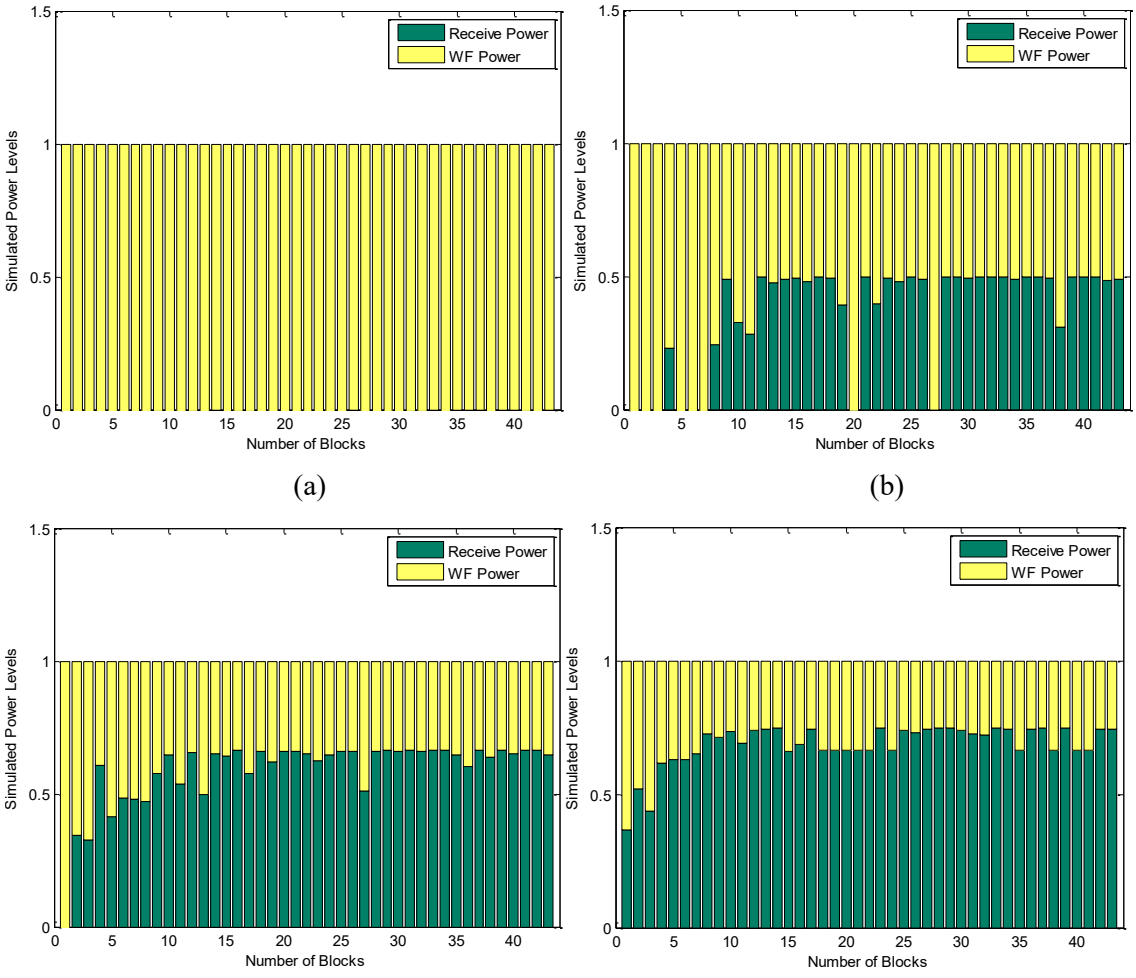


Fig. 5.13 Adaptive WF in MIMO channels (a) 1×1 (b) 2×2 (c) 3×3 (d) 4×4.

The third and last simulation exercise is aimed at the assessment of the rate adaptation capability, or the ASE, of the proposed scheme. Two targeted BER of 10^{-3} and 10^{-6} are pursued out of 1×1 and 2×2 MIMO channels of Rayleigh fading effect, while other settings are applied in the same above manner. In the beginning, the SNR mode switching operation needs to be defined properly. This can be done by inferring on the simulated BER versus SNR plot for single antennas as shown in Fig. 5.4. Following the style of Table 5.1 and counting for the specified BER targets, the segmented SNR switching ranges are extracted manually, the resulting algorithm of which is exerted as shown in Table 5.4 below. These values differ for other MIMO schemes, and as such the MRC effect can be applied by adding the received SNR values as discussed earlier in Chapter Three. The intention here is to show the power of space diversity and FET schemes, based on which the resulting ASE performance using the IMM-KF for CSI estimation and ideal informing link.

Table 5.4 Simulated SNR mode switching for 7 segments, single antennas, BER 10^{-3} and 10^{-6}.

Segment	γ_s range BER 10^{-3}	γ_s range BER 10^{-6}	Constellation
0	$0 \leq \gamma_s < 3$	$0 \leq \gamma_s < 9.4$	No Transmission
1	$3 \leq \gamma_s < 7.6$	$9.4 \leq \gamma_s < 13.1$	BPSK
2	$7.6 \leq \gamma_s < 11.5$	$13.1 \leq \gamma_s < 17.3$	QAM
3	$11.5 \leq \gamma_s < 15.75$	$17.3 \leq \gamma_s < 21.1$	8-QAM
4	$15.75 \leq \gamma_s < 17.9$	$21.1 \leq \gamma_s < 23.9$	16-QAM
5	$17.9 \leq \gamma_s < 21$	$23.9 \leq \gamma_s < 27.3$	32-QAM
6	$21 \leq \gamma_s < \infty$	$27.3 \leq \gamma_s < \infty$	64-QAM

The generated ASE curves that correspond to the switching algorithm of Table 5.4 are shown in Fig. 5.14. The ASE trends are obtained without varying the power for each achievable discrete rate, which could be employed to render these jump-style curves smoother. However, these jumps are favourable in the sense that instead of staying within one fixed-rate zone, the system may decide to switch to the next higher level. Upon sensing the receiver reporting is at a high enough SNR, which can comfortably occur without risks.

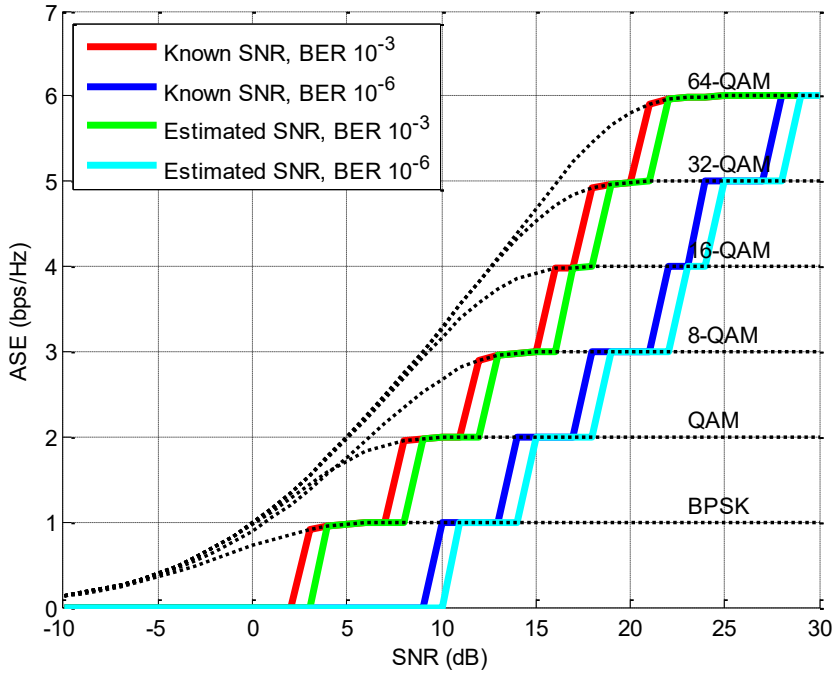


Fig. 5.14 ASE for MQAM signalling for 1 × 1 antennas, BER 10^{-3} and 10^{-6}.

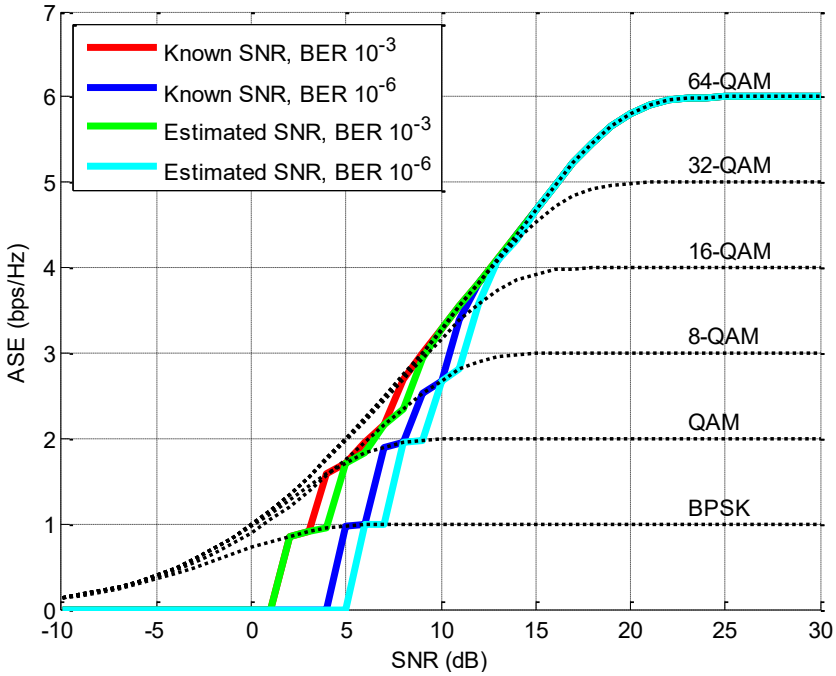


Fig. 5.15 ASE for MQAM signalling for 2 × 2 antennas, BER 10^{-3} and 10^{-6}.

Moreover, the ASE performances for the two targeted BERs of 10^{-3} and 10^{-6} are very distinct. As it is naturally anticipated, the latter BER target entails larger SNRs to be accomplished and hence the ASE traces are obviously seen displaced rightward toward such zones. However, still the ASE operation is capable of varying the

throughputs to the best extent possible and proportionate to the SNR and channel conditions in both scenarios.

The ASE performance boost caused by applying the 2×2 MIMO scheme is illustrated in Fig. 5.15. It is remarkably enough to get a quick look on this figure to realise how much the benefit is grasped by increasing the number of antennas. The overall ASE performance is greatly enhanced for both BER cases and no matter whether the SNR was actually known or estimated. Quite interesting is to see that the gap between the ASE performances are shrinking and shifting leftmost toward the smaller SNR zones. As expected, this is the main benefit gained by applying the space diversity of higher-order MIMO schemes. Even the differences between the known channels and estimated channels are getting infinitesimal now. Moreover, almost all traces are eventually getting a decent alignment with the optimal performance curves of each individual discrete constellation.

A final note in this regard, the stimulus of TCM and BICM schemes is not furtherly stretched here and the results of Fig. 5.7 and 5.9 are believed sufficient to infer on their performances. Any ACM scheme, whether conventional or newly proposed, merely decides on the preferable type of channel coding policy upon monitoring the instantaneous SNR and CSI estimate conditions as being informed back from the receiver. The TCM is typically selected in AWGN channels without severe fading, whereas the BICM is most favourable in case of adverse channel effects.

5.7 Summary

This chapter has dealt with strategic ACM schemes that constitute one of the prime pillars of the legacy wireless communication systems and the CR systems as well. Employing ACM has the potential to offer performance gains relative to fixed, uniform modulation. Three DoFs have been investigated, whereby the constellation rate, transmit power and channel coding are adapted with respect to the state of the channel. Generally, the complexity order of an optimal ACM scheme is prohibitive for real time applications that may run on portable devices with limited processing power and memory. Therefore, a suboptimal approach could represent an inevitable solution to many practical situations without any impact on the overall ACM performance.

Therefore, the ACM operation based on switching mode seemed to be a justifiable approach in reality and was widely accepted in diverse studies.

Under a FSMC channel assumption, a novel ACM exploiting the IMM-KF assembly has been proposed. This new scheme, which features the FET structure pronounced earlier in this work, proved to be appealing in terms of supplying sufficiently accurate CSI estimates. As such, the transmitter would have the capability to adhere to the varying channel conditions and hence dynamically adapt its transmission parameters accordingly.

The new paradigm IMM-KF scheme for ACM applications is well anticipated to evoke new horizons of unexplored applications, and can be used to reassess classical systems straightforwardly. This new ACM scheme has great potentials in pushing up the QoS and spectral efficiency levels, and hence qualifies to sustain further explorations in the future.

Chapter 6

AUTOMATIC MODULATION CLASSIFICATION

6.1 Introduction

AMC represents a prime and intricate task in modern intelligent radios. It shares with ACM and SS the same deep roots in the field of signal detection and estimation [143, 201]. The impetus of AMC was firstly catered in the military field a few decades ago. That may include tactical operations of electronic warfare, surveillance and threat analysis among others. Such operations are typically applied on an adversary intercepted signal to identify modulation patterns, jamming, counter measures and recovery. Soon after, such a prominent tool found its way towards civilian applications in the early 21st century when massive advancements took place in almost all technological domains. Stirred by the rising ubiquity demands in wireless communication systems, the industry discovered that AMC in conjunction with ACM and SS can deliver irresistible solutions for viable spectrum recycling.

With the provision of sufficient statistics, AMC effectively helps identify the various modulation forms of unknown noisy signals by using the MHT methodology. As a pre-processing stage located behind antennas, the AMC task is challenging especially in unfriendly or non-cooperative environments where many constellations and complex channel conditions are unidentified to the receiver. There are no particular templates available to match the intercepted observations seamlessly.

This chapter outlines the key characteristics of AMC that are centred on the LB method and also explores the new research trends in this manner. Given the many powerful features of the IMM-KF based on the FET scheme, such a paradigm is merited alongside the EM algorithm for efficient CSI estimation. This helps achieve seamless constellation recognition without extra computational loads compared to the classical MLE approach.

6.2 General Framework

The design of AMC essentially encompasses two stages; signal pre-processing and proper selection of the classification algorithm. The pre-processing tasks may include, but not be limited to, noise reduction, estimation, equalisation and power control, etc. Depending on the classification algorithm chosen in the second stage, pre-processing tasks with different levels of accuracy are required; some classification methods require precise estimates, whereas others are less sensitive to the unknown signal parameters.

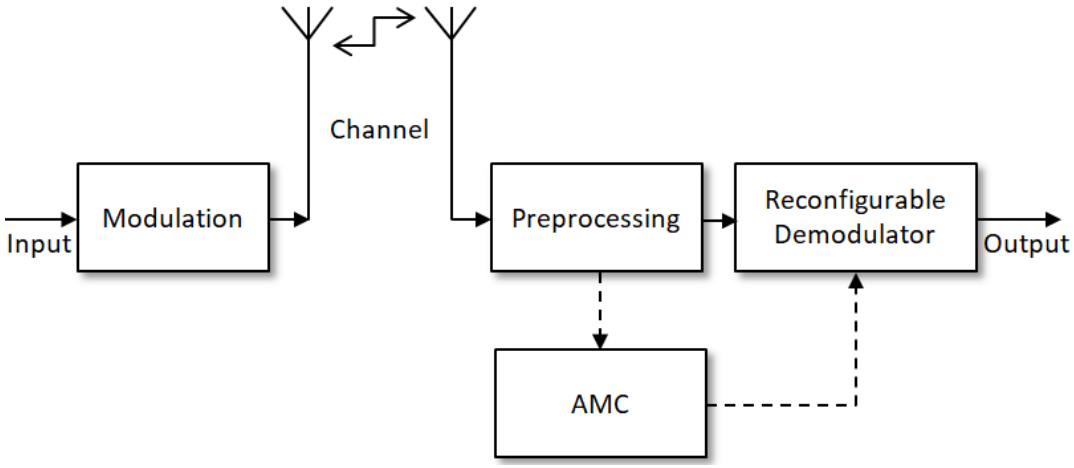


Fig. 6.1 Generic AMC system model.

A simplified block diagram of wireless CR that includes the AMC feature is shown in Fig. 6.1 [59, 61, 75], which can be easily generalised for any antennas combination. The AMC embedded in the CR is processed in tandem with a reconfigurable demodulator. A signal with unknown parameters is transmitted in RF, the modulation scheme of which is also unknown at the receiver side. This transmitted signal is impaired by the channel fluctuations and noise. The pre-processed baseband signal is furnished to the AMC unit, the modulation scheme of a transmitted data frame is recognised, and the software-configured demodulator is ordered to adjust the demodulation program accordingly. If the AMC processing is sufficiently fast, the software-configured demodulator can adapt itself to the scheme change and demodulate the incoming signal in real time manner.

There main criteria that determine the goodness aspects of an AMC are given below.

- Capability to classify as many modulation types as possible.
- Extension flexibility to accommodate new modulation schemes.
- High classification accuracy with respect to different SNR levels.
- Robust to various channel conditions, such as having a built-in CSI estimation and adaptation.
- Less complexity and computationally efficient.

Generally, a simple and fast classifier is highly desired. In practice, however, there is no such classifier that can meet all criteria. Therefore, the objective of this work is to develop a simple and forthright AMC strategy that is acumen to all criteria without major compromise as sensibly as possible. The fitness of the new AMC strategy is signified by the accurate identification of an important range of digital alphabets and the lesser complexity compared to other related methods. Owing to the suboptimal nature of the FB methods, this work therefore focuses solely on the LB approach, as it can generate optimal results in almost all situations. An AMC scheme that works blindly is also featured here.

6.3 Problem Statement

This section provides the AMC problem statement in terms of a linearly and digitally modulated signal transmitted over AWGN channel with Rayleigh fading effect. The most common digital modulation formats were discussed earlier in the context, but some of the terminologies are repeated here for clarity and completeness of presentation.

The received baseband signal that is being pre-processed and ready for classification using the MHT procedure is denoted as below [59, 60, 62, 64, 65, 68, 69, 74, 77, 78]

$$\mathcal{H}_i: \quad \boldsymbol{y} = h\boldsymbol{x}_i + \boldsymbol{n} \quad , \quad \boldsymbol{x}_i \in C_i^M \tag{6.1}$$

where $\boldsymbol{y} = [y(1)\ y(2) \dots y(K)]^{\mathrm{T}}$ is the complex-valued received signal observed over K samples, $h = ae^{j\theta}$ is the Rayleigh channel of a amplitude and θ phase variations,

$\boldsymbol{x} = [x(1)\ x(2) \ldots x(K)]^{\mathrm{T}}$ is the transmitted signal associated to constellation set mapping C_i, where the i^{th} hypothesis modulation type $\mathcal{H}_i$ could be assigned to any specific member of MPSK or MQAM family at a time, and $\boldsymbol{n} = [n(1)\ n(2)\ \ldots\ n(K)]^{\mathrm{T}}$ is the AWGN of zero-mean and unknown variance. It is assumed that the modulated symbols $\{x_i(k)\}_{i=1}^{M}$ are uniformly and independently distributed over C_i space. Other signal models based on passband forms are also available in the literature [76], but they all retain the same results except for minor deviations in the expressions.

Revising (6.1) in matrix form to consider all antennas' contributions yields

$$\mathcal{H}_i: \quad \boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X}_i + \boldsymbol{N} \ , \quad \boldsymbol{X}_i \in C_i^M \tag{6.2}$$

where $\boldsymbol{H}$ is the $N_t \times N_r$ channel matrix that is ZMCSCG and has *iid* entries. The above two models were given earlier in (3.2) and (3.3), but repeated for the purpose of complete exposition here. The task of AMC is to accurately decide on the best modulation type, from a pool of hypothesised known alphabets, which was actually sent by the transmitter.

6.4 Likelihood-Based Classifiers

AMC based on the LB statistical approach has become more popular and widely spread among other classification campaigns. When the CSI parameters are perfectly known to the classifiers, the optimality of the LB approach has thus become the prime drive for the elevated interest in the LB classifiers to get better classification accuracy. Even when the CSI are fully or partially estimated, the LB attributions are still capable of producing satisfactory classification results. Generally, such CSI estimate parameters can be attended either by using classical methods, or most effectively the proposed IMM-KF scheme, as will be shortly be seen in this chapter. The typical procedure of AMC based on the LB approach commonly constitutes two major steps. Firstly, the LF is developed for each modulation pattern, which can be manipulated to reduce complexity and survive various work conditions as desirable. Secondly, the LF of different hypothesised modulation patterns are compared against each other and the classification decision on the best candidate can be concluded accordingly.

The analysis of LB classifiers is intrinsic to the premise of Bayes system. The eventual classifier is realised by conducting a MHT, whereby $\mathcal{H}_i$ is arbitrarily assigned to the i^{th} modulation type out of M possible potentials. The MHT is a composite methodical problem with strong links to the joint and conditional PDFs of unknown parameters. In the realm of multivariate Gaussian statistics, the joint PDF conditioned on the i^{th} hypothesis and the unknowns can be inferred as follows [64, 65, 72, 74-78, 86, 88, 146, 201]

$$p\left(\mathbf{y}\middle|\{x\}_{i=1}^{M},\boldsymbol{\Theta}\right)=\prod_{k=1}^{K}\frac{1}{\sqrt{2\pi}\sigma_n}exp\left(-\frac{\left|y_k-h_kx_{k,i}\right|^2}{2\sigma_n^2}\right) \tag{6.2}$$

where the vector $\boldsymbol{\Theta}=[h\ \ \sigma_n]^{\mathrm{T}}$ signifies the unknown nuisance parameters. This can be easily amended to MIMO settings as shown below [62, 202, 203]

$$p\left(\boldsymbol{Y}\middle|\{x\}_{i=1}^{M},\boldsymbol{\Theta}\right)=\prod_{k=1}^{K}\frac{1}{(\pi\sigma_n)^{N_r}}exp\left(-\frac{\|\boldsymbol{Y}-\boldsymbol{HX_i}\|^2}{2\sigma_n^2}\right) \tag{6.3}$$

The above PDFs are directly involved in the computation of the LF that is hypothesised between the incoming signal and a look-up table (LUT) of known alphabets, and the decision is made based on the maximum of this function. Assuming that the likelihoods of all modulation formats are equally distributed, the resulting algorithm is dubbed optimum ML classifier in the sense that it achieves the minimum probability of error classification as per the rules of Bayes' decision theory. The optimum ML classifier is equivalent to the minimum distance classifier, which is famed by the Neyman-Pearson (NP) classifier [142, 201] and defined by the following Euclidean metric [107]

$$d_{min,i}=\sqrt{\|\boldsymbol{y}-\boldsymbol{x}_i\|^2} \tag{6.4}$$

where the argument $\|.\|$ means the vector norm and applied over all possible alphabets.

Generally, and irrespective of whether the classifier is considered optimum or semi-optimum, solutions to the composite MHT task can be attempted by appraising the unknown parameters as RVs. There are four different ML classification scenarios that can be accordingly distinguished, namely; the ALRT, GLRT, HLRT, and the QLRT [59-62] and a brief probing into each follows.

- ALRT: This reveals the LF of $\boldsymbol{y}$ under $\mathcal{H}_i$ is attained by averaging the PDF over unknown signal constellations and nuisance parameters

$$\Lambda_i^{\mathrm{ALRT}}(\boldsymbol{y}) = \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{\Theta}}[p(\boldsymbol{y}|\boldsymbol{x}_i, \boldsymbol{\Theta})] \tag{6.5}$$

  where $\mathbb{E}_{\mathrm{x}}[.]$ and $\mathbb{E}_{\Theta}[.]$ are the expectations operators (averaging is performed over all possible sequences of M symbols corresponding to the i^{th} modulation and vector of unknown parameters). The above ALRT algorithm assumes the statistical characteristics of unknown random variables are either exactly defined or fitted to the best, which is a vastly tedious task in reality. It is therefore practically not feasible in most CR applications, especially under large constellation size, due to its highly nonlinear complexity and computational burdens. The ALRT merely signifies the upper-bound limit for other algorithms.

- GLRT: This algorithm is taken as equivalent to ALRT, providing that the MLE of the nuisance parameters is used and, therefore, the performance of which is unavoidably inferior. Applying the GLRT algorithm under nested constellations is virtually a difficult task to achieve without making major errors [59]. Therefore, and like it has always been, a tradeoff between design priorities can be sorted out in most situations.

- QLRT and HLRT: The averaging is conducted over the unknown symbols only in these algorithms, and eventually the suitable estimate $\widehat{\boldsymbol{\Theta}}$ replaces $\boldsymbol{\Theta}$ as follows

$$\Lambda_i^{(\mathrm{Q})\mathrm{HLRT}}(\boldsymbol{y}) = \mathbb{E}_{\boldsymbol{x}}\left[p\left(\boldsymbol{y}, \widehat{\boldsymbol{\Theta}}|\boldsymbol{x}_i\right)\right] \tag{6.6}$$

  The MLE for sought parameters is at the heart of HLRT, while the QLRT may use other non-MLE tools. The attitude of NDA suggests that both algorithms can be realised blindly or semi-blindly, which is the case of this work.

Recall that the natural logarithm is monotonically increasing and Λ_i is positive-definite, the LLR defined as $\mathcal{L}_i = \mathcal{L}(\boldsymbol{y}/\mathcal{H}_i) = ln(\Lambda_i)$ is proved useful to simplify the AMC decision making. The decision criterion for binary hypothesis test (BHT) problem is hence constructed as the ratio between any two observations $\mathcal{H}_i$ and $\mathcal{H}_{i'}$ with prior probabilities P_i and $P_{i'}$, respectively, as given below

$$ln\frac{\mathcal{L}(\mathbf{y}/\mathcal{H}_i)}{\mathcal{L}(\mathbf{y}/\mathcal{H}_{i'})} = \mathcal{L}_i - \mathcal{L}_{i'} \gtrless_{\mathcal{H}_{i'}}^{\mathcal{H}_i} ln\frac{P_i}{P_{i'}} \tag{6.7}$$

The rightmost term in (6.7) is the threshold and its value equals 0 for equally likely modulation types, i.e., $P_i = P_{i'}$. This BHT can be uniformly generalised overall applicable M hypotheses space to yield the following detection rule

$$\hat{C}_i = arg_{i\in\{1,2,\ldots,M\}}\, max(\mathcal{L}_i) \tag{6.8}$$

The AMC algorithm is mainly tasked to maximise the LLR functions comprised in MHTs and decide the best modulation candidate with fewer errors as possible. The classifier performance can be quantified in terms of the individual probability of correct classification P_{CC}, which is the probability under hypothesis $\mathcal{H}_i$ given that modulation i is the accurate modulation

$$P_{CC} = \Pr(\text{correct}|\mathcal{H}_i) = \Pr[\mathcal{L}_i > \mathcal{L}_{i'}],\ \ \forall i \neq i' \tag{6.9}$$

The overall average probability of correct classification is not a favourable performance indicator since the individual probabilities cannot be assessed separately [76]. Therefore, the individual P_{CC} is considered in this work. The ML classifier for a typical M-ary MHT problem is illustrated in Fig. 6.2 [204, 205].

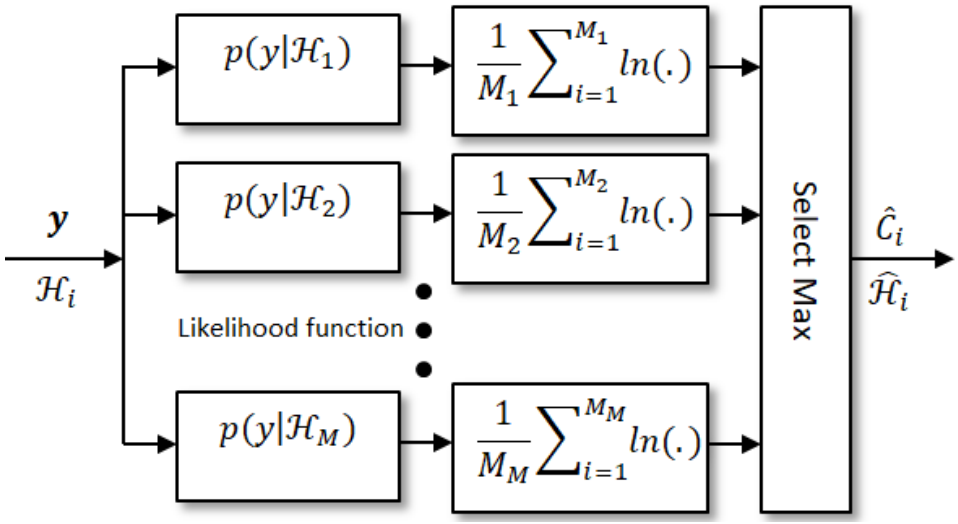


Fig. 6.2 ML classifier for M-ary constellations.

6.4.1 Coherent Scenario

This is an optimistic situation where the receiver collaborates with the transmitter for better signal retrieval [76]. In such circumstances, the coherent recovery of signal and perfectly known channel can be readily attained. Consequently, after the signal pre-processing, the coherent classifier has the full knowledge of all unknowns, except for the modulation type, where a flawless recognition by a consistent classifier is necessary. Empirically, the phase and timing attributes can be steadily tracked by synchronisation devices. Such coherent signal detection is also known as a non-blind classification.

The ALRT makes a good candidate for use under the coherence of hand shaking between transceivers. After dropping irrelevant constant terms from (6.2), (6.5) and (6.7) yields the following LLR for ALRT [59, 74, 76, 78]

$$\mathcal{L}_i^{ALRT}(\mathbf{y}) = \sum_{k=1}^{K} ln\left\{\frac{1}{M_i}\sum_{i=1}^{M_i} exp\left\{\frac{\left|y_k - ae^{j\theta}x_{k,i}\right|^2}{2\sigma_n^2}\right\}\right\} \tag{6.10}$$

and this will facilitate the decision making on the modulation type. In the case of BPSK and QPSK, the following closed-form expressions are given [74]

$$\mathcal{L}_{BPSK,i}^{ALRT}(\mathbf{y}) = ln\left\{\prod_{k=1}^{K}\frac{1}{2\pi\sigma_n^2} exp\left\{-\frac{(|y_k|^2 + a^2)}{2\sigma_n^2}\right\} \times cosh\left(\frac{2ay_{I,k}}{\sigma_n^2}\right)\right\} \tag{6.11}$$

$$\mathcal{L}_{QPSK,i}^{ALRT}(\mathbf{y})$$
$$= ln\left\{\prod_{k=1}^{K}\frac{1}{2\pi\sigma_n^2} exp\left\{-\frac{(|y_k|^2 + a^2)}{2\sigma_n^2}\right\} cosh\left(\frac{\sqrt{2}ay_{I,k}}{\sigma_n^2}\right) cosh\left(\frac{\sqrt{2}ay_{Q,k}}{\sigma_n^2}\right)\right\} \tag{6.12}$$

While for the MQAM, the amplitude of the received signal has Rician statistics and the LF of which is denoted by [76]

$$\mathcal{L}_{MQAM,i}^{ALRT}(\mathbf{y}) = ln\left\{\prod_{k=1}^{K}\frac{1}{M_i}\sum_{i}^{M_i}\frac{y_k}{\sigma_n^2} I_0\left(\frac{|y_k||x_{i,k}|}{\sigma_n^2}\right) \times exp\left(\frac{|y_k|^2 + |x_{i,k}|^2}{2\sigma_n^2}\right)\right\} \tag{6.13}$$

where I_0 is the modified Bessel function of the first kind with zero order as defined below

$$I_0(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} exp\{x\, cos(z)\}dz \tag{6.14}$$

When the number of symbols inside each data block is considerably high, i.e., assuming large value of the input sequence $K \to \infty$, then the following P_{CC} can be formulated [76]

$$P_{CC}(\mathbf{y}|\mathcal{H}_i) = \int_{-\infty}^{-\sqrt{K}\mu_{\mathrm{L}_i}} \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{L}_i}} e^{-\frac{v^2}{2\sigma_{\mathcal{L}_i}^2}}\, dv \tag{6.15}$$

where the mean and variance of $\mathcal{L}_i$ are defined below [73, 76], respectively

$$\mu_{\mathcal{L}_i} = \mathbb{E}[\mathcal{L}_i - \mathcal{L}_{i'}] \tag{6.16}$$

$$\sigma_{\mathcal{L}_i}^2 = \mathbb{E}[\mathcal{L}_i - \mathcal{L}_{i'}]^2 - \mu_{\mathcal{L}_i}^2 \tag{6.17}$$

The problem now reduces to the evaluation of the mean and variance values by the first and second order statistic of the random variable $\mathcal{L}_i$, or Λ_i, under each individual hypothesis and this would entail extensive numerical computations that are practically not feasible. It is important to realise the fact that despite the situation of applying the ALRT is idealistic, however, the classification results of which represent the upper-bound limits and hence can be usefully used to compare with other methods.

6.4.2 Non-coherent Scenario

Non-coherent detection, also known as blind, training-based or DA modulation classification, is very likely in diverse practical settings [59, 60, 62, 65, 68-77, 202, 203]. Of particular interest are the non-cooperative systems such as military surveillance, SS and some CRN applications, where the receiver has little or no knowledge of the transmitted signal. Other unknown parameters such as propagation delay, timing error, asynchronous phase and rapidly fluctuating channel, are likely not to be available under non-cooperative work conditions. This accordingly will render the statistics governing key variables uncertain over time and achieving the optimality in the detection process a challenging task. When one or more parameters are not explicitly acknowledged, a non-coherent estimator is unavoidable. There are two feasible approaches to alleviate the requirement for exact knowledge of these unavailable parameters. The first approach is to instantaneously estimate the parameters

of interest before admitting to the LRT functions in case their nature is of deterministic type. Otherwise, some of these parameters can also be viewed as random variables with known statistics and, hence, the realisation of the LFs can be attained by interrogating their joint PDFs as desirable.

In the quest of Q(H)LRT, the unknowns are treated as deterministic or random variables and replaced by their MLE or non-MLE for HLRT and QLRT, respectively. While, as stated above in the context, the unknown constellations are averaged over their known conditional PDFs. Therefore, (6.10) is revised to consider the estimated parameters as below

$$\mathcal{L}_i^{Q(H)LRT}(\mathbf{y}) = \sum_{k=1}^{K} ln\left\{\frac{1}{M_i}\sum_{m=1}^{M_i} exp\left\{\frac{\left|y_k - \hat{a}e^{j\hat{\theta}}x_{k,m,i}\right|^2}{2\hat{\sigma}_n^2}\right\}\right\} \tag{6.18}$$

The above statistical model constitutes approximations to the complex Bessel and hyperbolic mathematical functions and, hence, will generate a semi-optimal classifier based on the HLRT approach. The Bessel function is binding to the Rician PDF generated by the complex envelope of the received signal and noise of a Rayleigh distribution [206]. Using the in-phase and quadrature components, given earlier in the context, yields the following expression that can be obtained for the particular case of MPSK and MQAM [206]

$$\mathcal{L}_{MPSK,MQAM,i}^{Q(H)LRT}(\mathbf{y}) = \left|\sum_{k=1}^{K} ln\left\{\frac{1}{M_i}\sum_{m=1}^{M_i} exp\left\{\frac{\left(y_{I,k} - y_{Q,k}\right)^2}{2\sigma_n^2}\right\}\right\}\right|^{1/2} \tag{6.19}$$

Owing to the monotonic nature of the zero Bessel function, it has been shown that after furtherly simplifying the above, sufficient test statistics can be achieved [59, 76, 206, 207]

$$\begin{aligned}\mathcal{L}_{MPSK,MQAM,i}^{HLRT} - \mathcal{L}_{MPSK,MQAM,i'}^{HLRT} &= \left|\sum_{k=1}^{K}(y_k)^{M_i}\right| \\ &= \left[\left(\sum_{k=1}^{K}(y_{I,k})^{M_i}\right)^2 + \left(\sum_{k=1}^{K}(y_{Q,k})^{M_i}\right)^2\right]^{1/2} \underset{<}{\overset{>}{\mathcal{H}_{i'}}} \eta\end{aligned} \tag{6.20}$$

Without any arithmetic approximation, the threshold η is zero if all the probable modulations are equally likely, i.e., $P_i = P_{i'}$. The analytical closed form expressions for the MLE of the unknown signal amplitude, noise power, and phase, respectively, are given as follows [74], and partially [76, 78]

$$\hat{a}_i = \mathrm{Re}\left\{\sqrt{(\hat{\boldsymbol{x}}_i^H \boldsymbol{y})(\boldsymbol{y}^H \hat{\boldsymbol{x}}_i)}\right\} / \|\hat{\boldsymbol{x}}_i\|^2 \tag{6.21}$$

$$\hat{\sigma}_i = N^{-1}(\|\hat{x}_i\|^2 - |y^H \hat{x}_i|^2)/\|\hat{x}_i\|^2 \tag{6.22}$$

$$\hat{\theta}_i = \frac{-Im}{2} ln\left(\frac{\hat{\boldsymbol{x}}_i^H \boldsymbol{y}}{\boldsymbol{y}^H \hat{\boldsymbol{x}}_i}\right) \tag{6.23}$$

Further simplification of the above expressions can also be attended using the sequence length as below and without bias leakage control consideration [74, 107, 194]

$$\hat{a}_i = \left[\sum_{k=1}^{K} \frac{\left(y_{I,k}\hat{x}_{I,k,i} + y_{Q,k}\hat{x}_{Q,k,i}\right)}{\hat{x}_{k,i}^2}\right]^2 \tag{6.24}$$

$$\hat{\sigma}_i = K^{-1}\left(\sum_{k=1}^{N} y_{k,i}^2 - \hat{a}_i \hat{x}_{k,i}^2\right) \tag{6.25}$$

$$\hat{\theta}_i = {M_i}^{-1} tan^{-1}\left(\sum_{k=1}^{K}\left(\frac{y_{I,k}}{y_{Q,k}}\right)^{M_i}\right) \tag{6.26}$$

After careful examining the above MLE parameters, it can be easily deduced that they are a good match with their counterparts given earlier in Chapter Five; specifically those that were previously developed for the SNR estimation by Kerr and Gagliardi [197, 198].

The detected signal can be processed by applying (6.8) on the constellation space C_i. The decision rules of (6.8) and (6.20) can now have the reduced complex envelop MF, or so called the correlation receiver [107]

$$\hat{C}_i = \hat{x}_i = arg_{i\in\{1,2,\ldots,M\}} max|\boldsymbol{y}.\boldsymbol{C}| \tag{6.27}$$

where the sufficient test statistic in this case is equivalent to that given in (6.20). The phase estimate given in (6.26) can be used to adjust for adequate timing and, hence,

allow the optimum decision processor to pick the maximum SNR value. In this work, the approach of LUT is stressed by storing the known constellation sets on the receiver side memory. The above given estimates and the proposed LUT schematic are portrayed in Fig. 6.3, as a revisit of Fig. 6.2. Note that a schematic similar to the one in Fig. 6.3 was not accessible elsewhere, and made available here for the benefit of the explicit exposition.

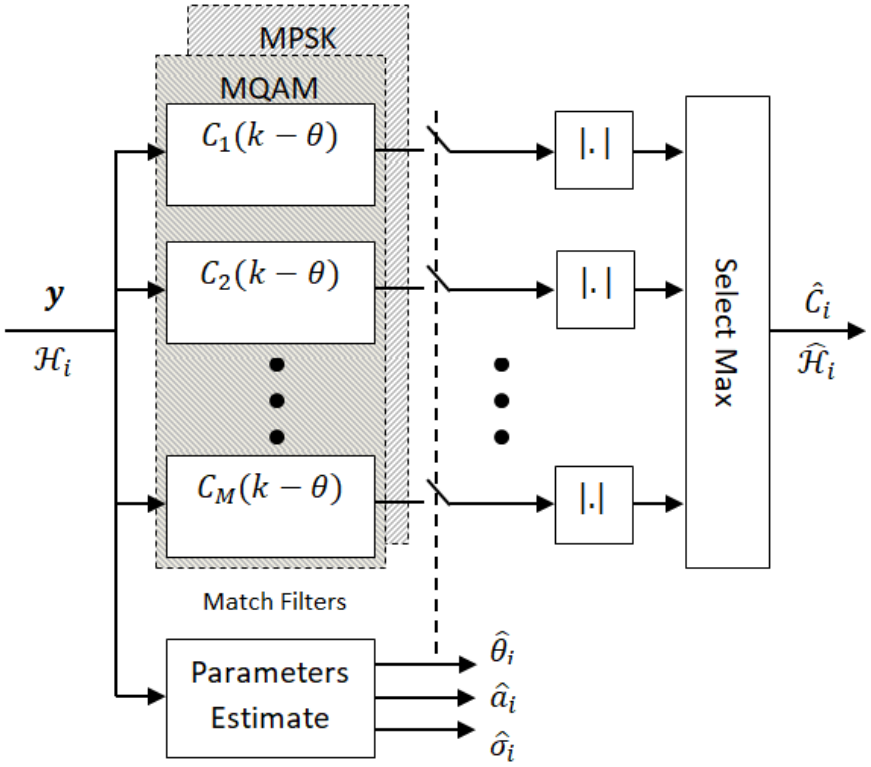


Fig. 6.3 Proposed illustration of MLE-LUT classification of M-ary Constellations.

On different occasions, such LUT arrangement is called the dictionary, where in which all possible modulations that can occur are assumed to be faithfully given and fixed [208]. Given that AMC is typically performed blindly in an unfriendly environment, it would be more beneficial to build a comprehensive dictionary that includes all possible candidate modulations employed by the extraneous target systems. Every time an unknown modulation is revealed, it needs to be added to this dictionary in order to enhance the chances of getting accurate classification outcomes in all situations. In the same manner, the LUT can also be called the library where all known modulations are safely stored. Upon the discovery of a new modulation, it can be

simply added to the existing library. Such a concept, where a semi-blind AMC module was formed using the minimum distance approach between the known and intercepted modulation, was proposed in [209]. Not to forget to mention as well that a finite set of quantised LFs of the known modulations, computed and stored in a LUT for comparison with their intercepted counterparts, was suggested in [73, 75]. Out of some others, only the above has been cited due to space here.

6.5 Joint Channel Estimation and Data Detection

Capitalizing on the above, it has become apparent the requirement to promote a joint link between the AMC and CSI estimation strategies. In the strict sense of AMC framework, the awareness of such joint channel estimation and data detection, dubbed here (JCEDD), is well established. The unknown CSI parameters that are accurately and seamlessly estimated would decisively contribute towards the successful operation of the AMC systems. Given the emanation of such quests are mostly encountered in crucially uncertain situations, the AMC empowered by CSI estimate should be acumen to endure adverse perplexes. The structure of which also needs to be less bulky in order to afford real-time processing as practically feasible. Therefore, this section aims to shed more light into the JCEDD approach along with supplementary options, which have been proved to be useful for the AMC and CSI estimation integrated performance.

Before moving with the JCEDD details, there is one thing worth to mention. When the LFs are aimed for signal detection using MIMO schemes, where SM is an essence, the first blind step seeks a BSS salvation [60, 209, 210]. The BSS based on ICA algorithm performs the pre-estimate of the channel matrix and sequence symbols, which may implicate the joint approximate diagonalization of eigenmatrices (JADE) approach on the road [60, 210]. It is also known that the ICA inherits random phase ambiguities. Therefore, the second step of LFs execution involves estimating these ambiguities for removal, where different methods such as the fourth-centroid [209] or the MoM [210] can be employed. Despite such estimators might look tempting for their simplicity, however, there is a possibility of getting near-ML performance, depending on the estimation problem at hand.

The research on JCEDD approach in the framework of blind AMC applications is extensively rich, and only a few visible studies are mentioned herein [211-220]. Given the fact that the wireless communications and signal processing industries have recently witnessed an explosive growth; the interest in blind techniques has also been expanded progressively. This naturally can be attributed to their multifaceted potentials in such important industrial sectors. A good overview of these blind techniques is given in [211].

Generally, the blind techniques can be branched into two main streams, either deterministic or statistical, depending on how the input signal is modelled. If the modelling assumes the input is a random variable with predefined statistics, the corresponding blind technique is to be statistical. On the other hand, if the modelling does not have any a priori description of the input random variable, or the statistics of which available but are not exploited, does not have a statistical description, the corresponding blind technique algorithm is called deterministic. The underlying algorithms of various existing blind techniques can also be furtherly classified into the MoM and the ML methods, and both can be employed alongside the above given two main streams [211].

The JCEDD can be applied to any kind of unknown parameters of interest, which evolve around the core parameter represented by the alphabet symbols. Moreover, the JCEDD algorithm can be approached via iteratively exchanging information between the CSI estimator and the signal detector, noting that the detection and classification words can be used interchangeably in the literature and in this work as well. The information exchange process is decomposed into two optimisation loops, the outer and the inner. The outer loop is a global optimisation algorithm, which searches for an optimal CSI estimate, while the inner loop is a ML algorithm that identifies the transmitted symbols [211-220]. These loops are interweaved, and sometimes called the upper level and the lower level, in tandem.

In other studies, for example, the repeated weighted boosting search algorithm was used for CSI estimate at the outer loop, which searches the MIMO channel space by evolving a population of MIMO channel matrices [212]. While at the inner loop, the optimised hierarchy reduced search algorithm aided detector, which is an advanced extension of the complex sphere decoder, was used [212]. The signal identification by

classifying both the modulation type and the STBC scheme in MIMO systems was considered as a joint classification problem [213]. The JCEDD was also extended to MIMO-OFDM systems in the framework of sparse Bayesian learning techniques [214]. It was also applied to the unknown co-scheduled user's modulation constellation size as part of the MU-MIMO detector [215]. Similarly, the size of MIMO channel matrix was tackled by exploiting the JCEDD, in order to improve the presumptions that inherently imply the knowledge of a priori information on the number of transmit antennas [216]. The per-survivor processing of best optimisation path using the VA algorithm as an effective trellis search engine in the context of JCEDD was elaborated [217-219]. The KF algorithm for CSI estimation was employed [214, 218, 220], whereas the GA procedure was implemented [219]. The systolic array structures based on lower diagonalization scheme was studied to enhance the KF computational speed in real-time practices [220].

On the side of design tools necessary for the JCEDD algorithm, the first step usually begins by tailoring a minimisation cost function, which should be suitably assigned to reflect on all the parameters of concern. Considering the MIMO configuration, this can be done by applying the MLE on the transmitted symbols $\boldsymbol{X_t}$ and the channel matrix $\boldsymbol{H}$ and then maximising the joint conditional PDF over $\boldsymbol{X}$ and $\boldsymbol{H}$ together. Alternatively, this can be achieved by minimising the following joint cost function [211, 212, 217-219]

$$\mathcal{F}\left(\breve{\boldsymbol{X}}, \breve{\boldsymbol{H}}\right) = \sum_{k=1}^{K} \|\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X}\|^2 \tag{6.28}$$

where all the constants have been ignored as they will cancel each other while calculating LRTs. Namely, the joint ML for channel and data estimation is obtained as below

$$\left(\widehat{\boldsymbol{X}}, \widehat{\boldsymbol{H}}\right) = arg\{min_{\breve{\boldsymbol{X}}, \breve{\boldsymbol{H}}}[\mathcal{F}\left(\breve{\boldsymbol{X}}, \breve{\boldsymbol{H}}\right)]\} \tag{6.29}$$

As can be seen from (6.29), the search for the optimal joint ML result is over the entire discrete space of all the possible transmitted alphabets and the MIMO channel matrix, which is computationally prohibitive in reality. Therefore, the complexity of such optimisation process needs to be reduced to tractable levels. This can be done if each single search cycle is decomposed into two iterative loops, the first over all the

possible data symbols and then over all the channel matrix entries. This can be formed as below

$$(\widehat{\boldsymbol{X}}, \widehat{\boldsymbol{H}}) = arg\left\{min_{\breve{\boldsymbol{H}}}\{min_{\breve{\boldsymbol{X}}}[\mathcal{F}(\breve{\boldsymbol{X}}, \breve{\boldsymbol{H}})]\}\right\} \tag{6.30}$$

Such interlaced loops are repeated over the complete space of all data and channel entries in effect, and until a breakdown of the most suitable per-survivor paths are obtained with minimal errors as possible. Another efficient iteration tool is explored in what follows.

6.5.1 Expectation-Maximisation

The EM is widely adopted as an effective tool for iterative computation of the MLE cost function in the ACM framework [61, 62, 76, 77, 88, 203, 211, 217, 221-223]. This algorithm is very useful in problems where the joint use of observed data $\boldsymbol{Y}$ and hidden data $\boldsymbol{\Theta}$, or simply $\boldsymbol{X}$ and $\boldsymbol{H}$, or to be defined otherwise, retains further simplifications in the maximisation of the LFs. The observed data along with the hidden data constitute the complete data set $(\boldsymbol{X}, \boldsymbol{H})$ necessary for such maximisation process. Areas of application may include speech recognition and the estimation of mixture distribution parameters, among many others.

The EM algorithm comprises two steps, namely; the expectation step (E-step) and the maximisation step (M-step). In the E-step, the expectation of the LLF of the complete data given the observed data is evaluated. While in the M-step, new estimates of the unknowns are obtained by maximising the expectation computed in the E-step. The EM algorithm possesses the property of monotonically increasing at each step to produce the expected maximum of LFs when it converges. Other variants of the EM algorithm also exist. For example, the expectation/conditional maximisation algorithm is used to replace the complicated M-step by computationally simpler conditional maximisation stages, while the E-step remains unchanged, and the monotone convergence property of the EM algorithm is preserved [77, 221, 222]. In order to speed up the convergence of the EM algorithm, the method of K-means to find a good starting point for the EM algorithm and converge quickly was proposed [221]. In contrast to the standard EM algorithm, where an offline and batch process of the whole data at each iteration is assumed, an online EM version was investigated for the purpose

of alleviating the computational complexity as well as the data storage requirements [222]. To seek proper initialisation points, simulated annealing was investigated as part of a generalised EM method in the context of assisted classification using a multiple sensors environment [223]. However, a precaution needs to be exercised as the EM algorithm does not always guarantee convergence to the global maximum. That may occur when the LLF has multiple local maxima, with the point of convergence depending on the initial estimates.

Keeping the above in mind, the EM algorithm can now be explicably aspired. Suppose the only unknown parameter is $\boldsymbol{H}$, then the MLE is denoted as follows [77, 211, 221-223]

$$\widehat{\boldsymbol{H}} = arg\{max_{\boldsymbol{H}}\{ln\, p(\boldsymbol{Y}|\boldsymbol{H})\}\} = arg\left\{max_{\boldsymbol{H}}\left\{ln\left[\sum_{\boldsymbol{X}} ln\, p(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{H})p(\boldsymbol{X})\right]\right\}\right\} \tag{6.31}$$

From (6.31), it is obvious that there is no closed-form solution for $\widehat{\boldsymbol{H}}$ since $p(\boldsymbol{Y}|\boldsymbol{H})$ is of a mixed Gaussian distribution nature. A circumvent to this issue is to use the EM algorithm to iteratively compute the MLE by using two data sets. The first data set is incomplete and comprises the observations $\boldsymbol{Y}$, while the second is complete and is represented by $\boldsymbol{Z} = [\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}]^{\mathrm{T}}$, which is the observation and the unknown symbols.

The E-step is thereby attended as below

$$\mathcal{Q}\left(\boldsymbol{X}, \widehat{\boldsymbol{H}}^{k-1}\right) = \mathbb{E}_{\boldsymbol{Z}}\left\{ln\, p\left(\boldsymbol{Z}|\boldsymbol{H}|\boldsymbol{Y}, \widehat{\boldsymbol{H}}^{k-1}\right)\right\} = \mathbb{E}_{\boldsymbol{X}}\left\{p\left(\boldsymbol{X}|\boldsymbol{Y}, \widehat{\boldsymbol{H}}^{k-1}\right) ln\, p\left(\boldsymbol{Z}|\widehat{\boldsymbol{H}}^{k-1}\right)\right\} \tag{6.32}$$

where $p\left(\boldsymbol{X}|\boldsymbol{Y}, \widehat{\boldsymbol{H}}^{k-1}\right)$ is the a posteriori probability of the modulation symbol vector $\boldsymbol{X}$ conditioned on $\boldsymbol{Y}$ and $\widehat{\boldsymbol{H}}^{k-1}$, and k denotes the k^{th} iteration of estimation and detection.

The parameter $\widehat{\boldsymbol{H}}^{k}$ maximises $\mathcal{Q}$ in the M-step as follows

$$\widehat{\boldsymbol{H}}^{k} = argmax_{\boldsymbol{H}}\left\{\mathcal{Q}\left(\boldsymbol{X}, \widehat{\boldsymbol{H}}^{k-1}\right)\right\} \tag{6.33}$$

The M-step (6.33) signifies the most surviving path essential for the MLE algorithm, while the E-step (6.32) forms the basis to maximise (6.5) and (6.6) or minimise (6.28). Starting from an initial estimate $\widehat{\boldsymbol{H}}^{1}$, or namely $\widehat{\boldsymbol{H}}(1)$, the convergence of iterative EM algorithm to a stationary point of the LLR is guaranteed under mild regularity conditions.

6.5.2 Space Reduction

The scale of complexity load constitutes a continuous challenge in a wide variety of wireless communication systems over all the times. It has been considerably endeavoured to find suitable means of effective reductions to the best permissible extents. Specialised industrial and academic sectors saved no efforts to render such paramount goal feasible. Generally, the attitude of subspace detection, which is typically based on channel decomposition, offers good options to flexibly compromise between performance and complexity. Of several remedial ways to such a challenge is the SVD, which attractively retains a massive reduction in the size of MIMO systems and their computations. That is owed to the SVD strength in decomposing wireless channels into a few parallel paths. Chapter Three has more details in this regard.

There are a number of techniques aimed at achieving fair amounts of MIMO size lowering and tailored to fit in the MLE and CSI estimation environments. For example, some sorts of subspace structures have been discussed in the context of blind identification techniques [211]. The main idea of such structures is that the whole channel vector, or part of it, can line up in a one dimensional subspace of either the observation statistics or a block of noiseless observations. The subspace algorithms are attractive in the sense that the CSI estimates can often be obtained in closed form expressions, which are formed in quadratic cost functions for optimisation.

Another interesting space reduction algorithm called sphere decoding, also known as sphere packing, reduces the number of visited lattice points [45, 107, 171]. In the context of JCEDD, the sphere decoding has been studied to control the complexity of a receiver to achieve as minimum performance errors as possible [224]. However, the sphere decoding for ML solution was reliant on another matrix decomposition scheme called the QRD, which was highlighted earlier in the context of SVD approach in Chapter Four [160]. Moreover, and despite they can achieve ML performance, sphere detectors require substantially higher complexity [225]. Another interesting subspace form using the quadrature left decomposition (QLD) was discussed [220, 225], and further details on the QRD is also accessible [226]. Both these two studies suggested further space reduction by using the WRD approach, where W and R stand for matrices' names used in the analysis therein, which can be attended by swapping particular layers in the punctured structures [225, 226]. The complexity overheads were

also evaluated and major savings were claimed while achieving near optimum classification performances.

Admitting all the given above, the topic of computation and complexity provisions in the real world of wireless communication systems is actually quite pressing. In the awareness of this undeniable fact, the CSI estimation scheme based on the new paradigm of FET and suitably customised to allow an efficient IMM-KF application drives gracefully into this course. All such blended ingredients can achieve satisfactory classification results, and above all, with lesser complexity than the aforementioned methods. This is detailed in the following section.

6.6 Proposed AMC System

Resilient CSI estimation plays a key role in the overall AMC systems' performance, especially of higher-order modulations, and KFs are very attractive in such domains. The use of KFs in the framework of AMC applications is not new and all studies concurred on their prudent behaviour [211, 214, 217, 220]. Combined with EM algorithms, the robust KF structure suits well the JCEDD environment. More complexity drop and reliable JCEDD enactment can be realised by employing the embedded IMM-KF scheme. This new paradigm is planned to have an effective CSI estimate supplied for the ML to action signal detection flawlessly.

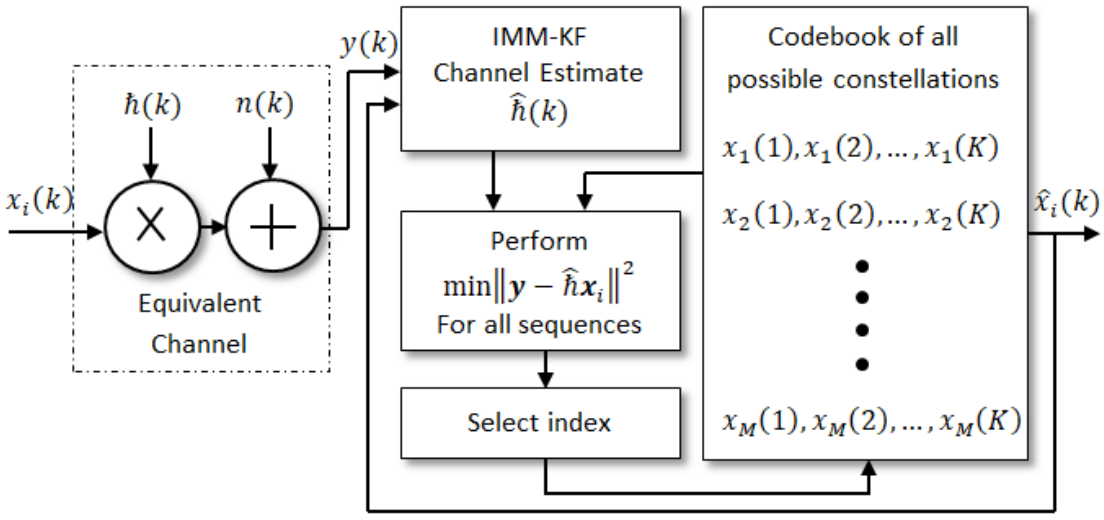


Fig. 6.4 Simplified AMC system diagram based on IMM-KF for CSI estimation.

Featuring the AR(1) Gauss-Markov signal model, which invokes the Jake and Clark attributes, and the FET observation model for IMM-KF, EM cycles for the ML signal detection can hence be developed. All these modelling procedures were discussed in Chapter Four. Guided by the system diagram in [218], among others, the proposed ACM system based on the IMM-KF for efficient CSI estimation is depicted in Fig. 6.4 above. The overlapped inner and outer loops for the EM iterative computation to facilitate the JCEDD are quite obvious in this figure. The iteration between CSI estimator and data decoding carries on until convergence is obtained.

By referring to the IMM-KF algorithm described in Table 4.1, there is one exception related to the recursive computation of KF state estimate that needs to be signified. The KF gain and residual sequence must now consider the detected sequence symbol which belongs to one of the known candidate alphabets [149, 214]. Putting together all KF equations, for the sake of complete analysis, the predictions and estimates of channel state and covariance parameters can be attended as below.

The predictions for the system model as described earlier in (4.31) and (4.32) are below

$$\begin{cases} \widehat{\hbar}_{k|k-1} = \widehat{\hbar}_{k-1|k-1} \\ P_{k|k-1} = P_{k-1|k-1} + \sigma_n^2 \end{cases} \tag{6.34}$$

and the CSI estimate can be recursively updated as follows

$$\begin{cases} \widehat{\hbar}_{k|k} = \widehat{\hbar}_{k|k-1} + G_k z_k \\ G_k \triangleq P_{k|k-1} \hat{x}_{i,k-1}^{\mathrm{T}} {S_k}^{-1} \\ P_{k|k} = (1 - G_k) P_{k|k-1} \end{cases} \tag{6.35}$$

where the KF gain G_k has been affected by $\hat{x}_{i,k-1}$, which is a priori identified symbol as per hypothesis $\mathcal{H}_i$ and imitated by the residual sequence $z_k = \left(y_k - \hat{x}_{i,k-1}\widehat{\hbar}_{k|k-1}\right)$ and its covariance $S_k = \hat{x}_{i,k-1} P_{k-1|k-1} \hat{x}_{i,k-1}^{\mathrm{T}} + \sigma_n^2$. The rest of IMM reckoning remains intact.

The EM algorithm also needs to accommodate the estimated FET parameter as below.

- E-step

$$\mathcal{Q}\left(x, \widehat{\hbar}^{i-1}\right) = \mathbb{E}_{\boldsymbol{z}}\left\{ln\, p\left(\boldsymbol{z}|\hbar|\boldsymbol{y}, \widehat{\hbar}^{i-1}\right)\right\} = \mathbb{E}_{\boldsymbol{x}}\left\{p\left(\boldsymbol{x}|\boldsymbol{y}, \widehat{\hbar}^{i-1}\right) ln\, p\left(\boldsymbol{z}|\widehat{\hbar}^{i-1}\right)\right\} \tag{6.36}$$

- M-step

$$\widehat{\hbar}^{i} = argmax_{\hbar}\{\mathcal{Q}(x, \widehat{\hbar}^{i-1})\} \tag{6.37}$$

The most surviving path, which is essential for the MLE algorithm and signified in the M-step (6.37), is directly realised by the IMM-KF. The E-step (6.36) relates to both the IMM-KF and the MLE algorithms. So the EM procedure performs as follows. Given a priori known modulation symbol, apply the E-step to evaluate the metric that links this symbol with the previous CSI estimate, the M-step is then applied to maximise the metric that assesses the compatibility between the CSI estimate and the received one. The whole process is repeated over the entire ensemble of all known alphabets and sequence points, refer to Fig. 6.4 above. Such a search cycle is obviously tedious and prohibitive in reality, especially for high-order alphabets. Thus, any effort to reduce the burden of this exhaustive search engine would be highly valued in seamless systems.

6.6.1 Complexity Assessment

The complexity of any ACM scheme is proportional to the numbers of code sequence and alphabets, but since the sequence is intact, the focus here will be on the alphabets order only. Recall that the subspace is reduced by combining the MRC and SVD algorithms in the FET for IMM-KF, the complexity of proposed classifiers can be easily assessed as it is dropped from $\mathcal{O}(M^{N_t})$ [60, 226] to $\mathcal{O}(M)$ for each step in the EM algorithm. The optimal MLE requires $\mathcal{O}(M)$ for each iteration in (6.36) and (6.37), while the proposed IMM-KF needs one iteration in (6.36) and its estimation can be readily plugged into (6.37) without extra burdens. The figure of $\mathcal{O}(M)$ in reality can grow overly unfeasible for larger constellations, while the IMM-KF load is fixed and thereby converges easily. The IMM-KF is considerably cheaper than MLE and hence can be advocated for seamless ACM systems.

6.7 Simulation Results and Discussion

Simulations are conducted to evaluate the AMC performance using the proposed IMM-KF and compared to that of MLE and perfectly known channel. Similar to the exercises in previous chapters, a 3-KF bank is employed to facilitate the IMM algorithm

with initial probabilities of 0.1, 0.7 and 0.1 for each individual filter, while the transitional probabilities of 0.7, 0.1 and 0.1 represent the local, neighbour and remote FSMC states, respectively. Each KF has a single σ_n^2 value, such as 0.01, 0.1 and 1, to resemble mild, moderate and harsh channel conditions, respectively. The values for $\hat{h}_{0|0} > 0$ and $P_{1|0} \geq 1$ also can be arbitrarily initialised.

A challenging 3-tuple set of nesting constellations {BPSK, 8-PSK, 16-PSK} is generated using the celebrated procedure. The results of 10^3 Monte Carlo runs over an uncorrelated slow-fading Rayleigh channel and 1×1 and 2×2 antennas are shown in Figs. 6.5 and 6.6, respectively. The results illustrate correct classification P_{CC} calculated for the IMM-KF, MLE and known channel, where the latter serves as an upper-bound performance. Table 6.1 below shows their correct classification rates at 0 dB.

Table 6.1 Set-1 correct classification rates at 0 dB.

	1×1 antennas			**2×2 antennas**		
	Known channel	MLE	IMM-KF	Known channel	MLE	IMM-KF
BPSK	0.99	0.99	0.99	1	1	1
8-PSK	0.85	0.85	0.85	0.98	0.98	0.98
16-PSK	0.57	0.57	0.57	0.75	0.75	0.75

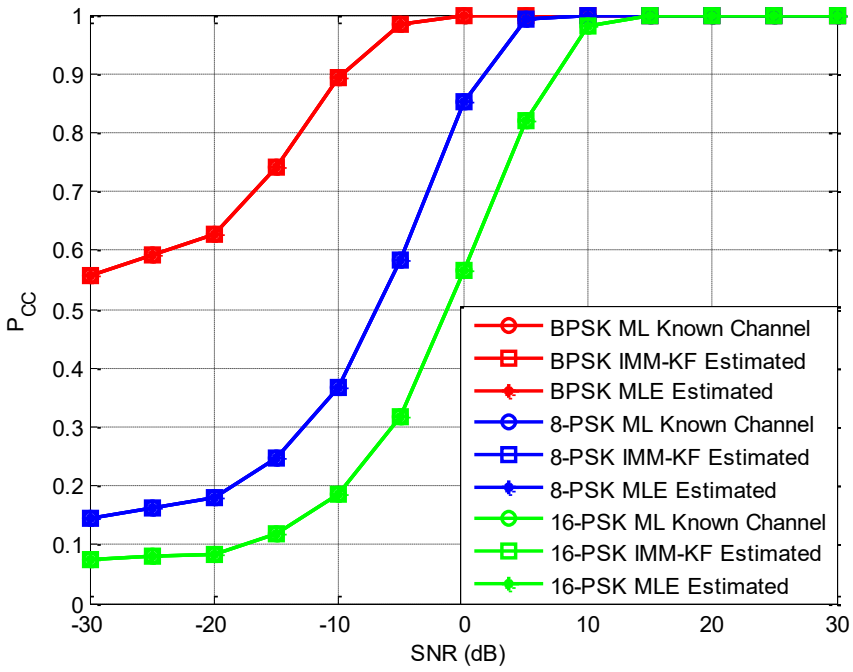


Fig. 6.5 P_{CC} for PSK constellations using 1×1 antennas.

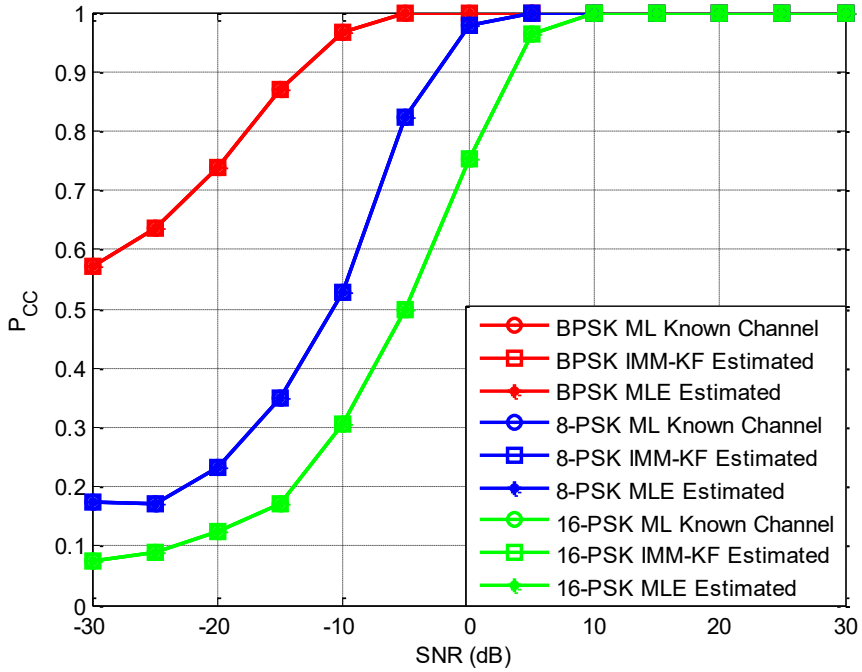


Fig. 6.6 P_{CC} for PSK constellations using 2×2 antennas.

The same simulation scenarios are repeated using another 3-tuple set of different constellations, namely {QPSK, 16-QAM, 64-QAM}, which is tailored to realistic common systems. The ensuing P_{CC} trends are depicted in Figs. 6.7 and 6.8, for and 1×1 and 2×2 antennas, respectively, and the values of which calculated at 0 dB given in Table 6.2 below.

Table 6.2 Set-2 correct classification rates at 0 dB.

	1×1 antennas			**2×2 antennas**		
	Known channel	MLE	IMM-KF	Known channel	MLE	IMM-KF
QPSK	0.75	0.75	0.75	0.99	0.99	0.99
16-QAM	0.27	0.27	0.27	0.67	0.67	0.67
64-QAM	0.04	0.04	0.04	0.30	0.30	0.30

The results show that the proposed IMM-KF performs similarly to the optimal MLE and both coincide with the known channel irrespective of antennas. The increased SNR yields for a gradual dominance of the signal covariance matrix and to spread its eigenvalues, which enhances the signal detection accordingly. All methods are also effective in combating the nesting effect to a certain limit, whereas the trends of correct classification clearly shifting down for larger constellations' sizes. A classifier to attain

correct decisions hence gets harder in the severity of lesser spaces between constellations' points. Such sternness is very evident in the cases of 16-PSK, 16-QAM, 64-QAM as the spaces between adjacent constellation points gets more stuffed and the detection hence becomes very compelling. For the same M-ary, the QAM suffers from magnitude and phase variants that are even harder to identify.

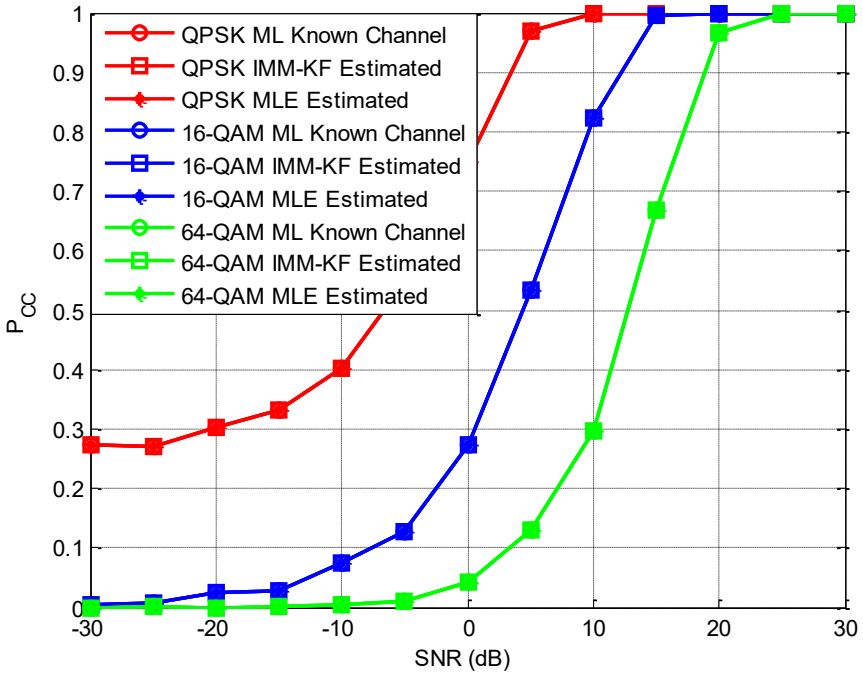


Fig. 6.7 P_{CC} for QPSK and QAM constellations using 1×1 antennas.

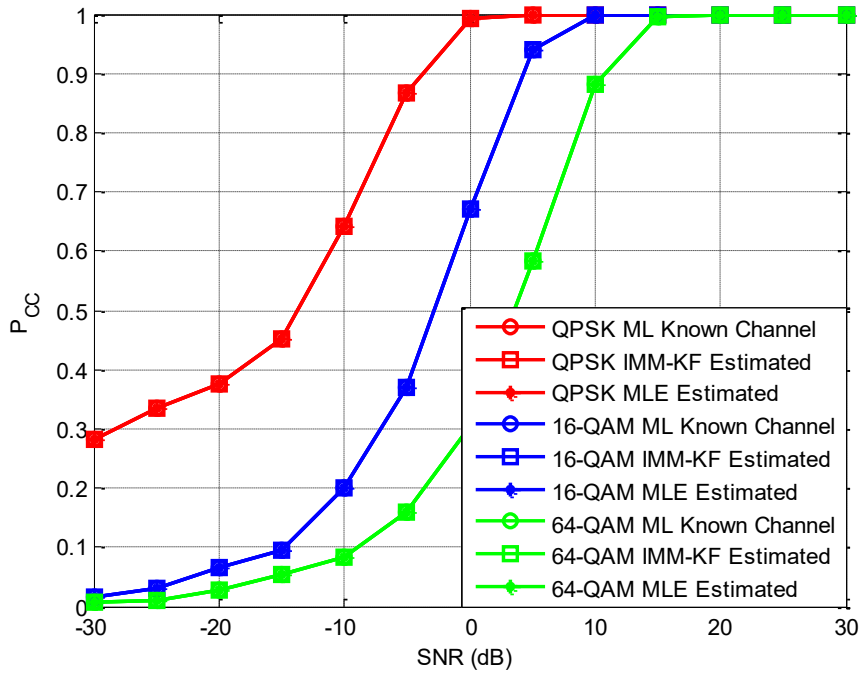


Fig. 6.8 P_{CC} for QPSK and QAM constellations using 2×2 antennas.

6.8 Summary

This chapter has presented the basics for AMC using the LB approach in preference over that of using the FB approach. That is mainly due the tendency of LB algorithms to produce optimum or near-optimum classification outcomes are highly likely compared to any other approach. Various LFs have also been illustrated, such as ALRT, GLRT and Q(H)LRT and the selection among which is highly dependent on how much statistical knowledge on the unknown parameters are available to hand. Such unknowns, whether treated as random deterministic or variables along with their statistical distributions, determines which algorithms fits well for particular applications.

This chapter also featured a novel AMC approach based on the renowned IMM-KF for CSI estimation and to support the ML classification function. The whole combination of which was admitted to the EM overlapped cycles to carry out the iterative computation of both the CSI estimation and the signal detection processes. An assessment of the underlying complexity and computational requirements revealed that the AMC based on IMM-KF has advantageously less overheads compared to the classical MLE for CSI estimation. Simulation results also revealed that classification errors of the newly proposed paradigm are almost compatible to those obtained when the channel parameters were perfectly known, in other words the same as of those obtained by the optimum MLE algorithm. Therefore, this chapter can come up with the prime conclusion of having the IMM-KF advocates better for seamless processing in practical AMC systems. This would render farsighted new opportunities in the area of efficient and resilient AMC systems, which were not perceived before, readily accessible.

Chapter 7

SPECTRUM ESTIMATION AND SENSING

7.1 Introduction

SS is the totem pole in the landscape of CR applications and around which all other components evolve. Like ACM and AMC, SS has deep roots in the statistical estimation theory. Two main classes of spectrum estimation techniques can be named; parametric and non-parametric. The first commonly relies on postulated models for the stochastic process under investigation, and this may include AR, AR moving-average (ARMA), and MA. This method typically provides high spectrum resolution without making any assumptions on the missing or unknown parameters. While the second does not require stochastic modelling, but assumptions regarding the unknown parameters are usually made. However, they possess poor consistency and biasing, which may lead to poor spectral resolution. As an enhancement, ample windowing options can be used to truncate the spectrum side lobes. Examples of non-parametric techniques are periodogram, MVUE and MTM. For the SS function to be complete, an appropriate detection policy is essential in addition to the spectrum estimation. The latter constitutes the BHT as a genuine decision making in most SS applications. The selection of SS elements and procedures is application specific and depends on the design constraints and requirements.

This chapter is devoted to the non-parametric SS approach with particular emphasis on the MTM due to its sturdy performance under various harsh environments. Featuring the IMM-KF structure, the adaptation of MTM parameters, namely; AT and OFD, is readily amenable as will be explored herein. There are several novel applications achieved in this chapter, but instead of being confined within one place, they are spread across the context and highlighted wherever they may occur.

7.2 General Framework

In CR systems and CRNs, an important part of signal processing is allocated for SS applications. It is primarily tasked to reliably and efficiently identify idle spectrum holes and share them among bidding users without side effects. The duty of SS is to decide if there is a signal at the receiver input or not, which is normally caused by a PU. No signal means a spectrum hole is available for potential SUs to use, and vice versa is true.

The awareness, analysis and decision making related to spectrum occupation and availability recently paved the way for a new research of DSA and dynamic spectrum management (DSM) [2, 55, 91, 92, 94, 95, 97, 98, 102, 103, 104]. This will enable the xG radio systems to utilise the spectrum more efficiently without interfering or harming PUs. Recall from Chapter Two, Fig. 2.7 reflects on the DSA action performed by a typical CR. Fig. 2.7 also demonstrates that the time duration, centre frequency and width of occupied spectrum holes are considered as random parameters. The carrier power of each occupied channel also varies depending on the spatiotemporal mobility of PUs. If particular channels used by SUs deteriorate or are reclaimed by PUs, SUs halt and try to find other available channels and resume their transmission.

The ON-OFF transmission in CRNs is termed as spectrum handoff [2, 227, 228]. Since transmissions of SUs are suspended during the phase of spectrum handoff, they will experience longer packet delay. Therefore, a good spectrum handoff mechanism should furnish SUs with smooth frequency shift with minimum latency. Interweaving or laying techniques could be sought as alternative options among other DSA models, or by using brain-inspired decentralised DSM based on self-organizing maps. Generally, SS is temporal when it decides either a PU is present or not (ON-OFF), while it is spatial given distant separations between PUs and SUs; all can thereby occupy the same bands concurrently if there is no restriction concerning the geographical coverage areas [93].

A useful tool called spectrogram, or spectrum waterfalls, can be theoretically examined. Such a tool has a visual of the random variations of frequency spectrum with respect to time. As shown in Fig. 7.1, typical PUs' activities are shown as peaks lasting for random durations over selective frequency channels. Any spectrum estimator strives

to allocate empty holes not being used and jump in between them. SUs may vacate their occupied channels upon PUs request within less than 2 seconds as per the IEEE 802.22 [8, 229]. Therefore, the objective of any SS technique is to randomly or successively check other channels for possible availability within the above given time or shorter. This is an actively ongoing study on sustainable spectrum management and will not be attempted here.

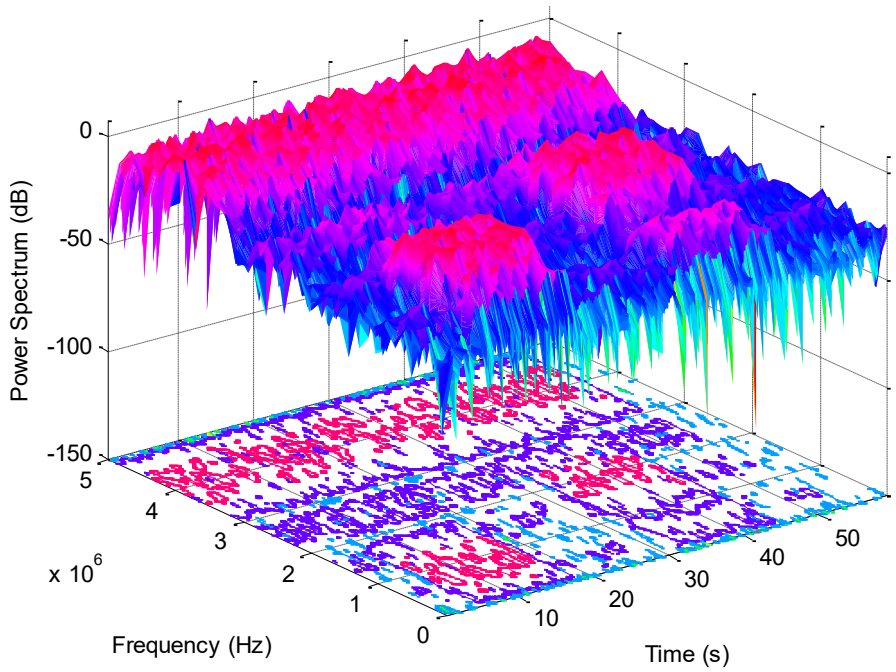


Fig. 7.1 A hypothetical spectrogram for frequency bands utilisation.

7.2.1 Sensing Scenario

As stated earlier in the context, the notion of CR has arisen over legacy wireless communication systems to deal with multiple radio access technologies aimed at spectrum sharing. By overlaying or underlaying, unlicensed SUs can fill in the empty bands without harming licensed PUs. The SS role hence institutes an essential part of the CR systems to dynamically detect and use the spectrum holes. A typical scenario of wireless communication system with CR perception is depicted in Fig. 7.2.

Licensed PUs are assumed to be randomly connecting to their service provider through a BS with certain territorial coverage, as in normal mobile systems. Certain frequencies are assigned within each cell and handed over to another carrier without

interruption during inter-cell transients. The pre-assigned number of channels is determined within each cell as per governing standards and based on the technology generation in place.

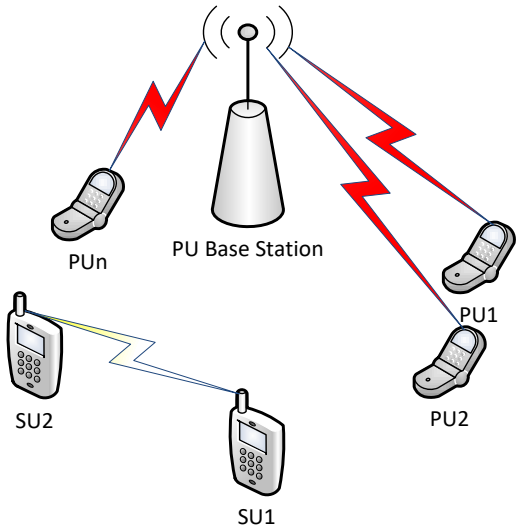


Fig. 7.2 Typical non-cooperative CR wireless communication scenario.

SS is a key enabler for CR systems and CRNs to quantify the spectrum awareness, user requirements and applications [91, 100]. When a certain band is detected as not being used by licensed PUs at a particular time in a particular geolocation, unlicensed SUs can utilise the band, i.e., there exists a spectrum opportunity. SS can thereby be performed in the time, frequency, and spatial domains [91, 93].

SS is generally tasked to get the knowledge on the spectrum occupancy status, which adheres to any of the following three different styles [95, 98]

- Geolocation and database.
- Listening to the cognitive pilot channel or PU beacons.
- Local at the CR terminal.

The FCC, at some stage, mandated to use the geolocation and database as the main spectrum awareness to facilitate the use of TVWS [95]. CR systems must hence download all the information about TVWS, such as; the RF environments, traffic patterns, locations, transmit powers, etc., from a remote knowledge database. Though it

is a database aided SS [98], however, it raised concerns about the free and efficient access of such remote knowledge database. Therefore, the most effective style is to identify the spectrum availability by detecting the transmission signals of PUs within the range of a CR. However, the direct intervention and inferring the channel between primary transmitters and receivers is difficult in reality. Therefore, most existing SS algorithms focus on the detection of the primary transmitted signal based on the local observations at the CR node. Aiming at the enhancement of detection probability, SS can be equipped with various options of reliable signal detection techniques, and the one that is selected in this work belongs to the non-parametric spectrum estimation category and consolidated with the BHT for decision making.

As it is not expected that all licensed PUs request connection most of the time, then opportunistic SUs are expected to randomly emerge to try to occupy unused channels. SUs have to be smart, listen carefully to the PUs activities and maintain SS to dwell in an empty white hole and establish desirable links. The central node of PUs, or BS, is non-cooperative in this scenario, as there is no spectrum information transmitted to SUs. The CR engines in the platforms of SUs can monitor as many smart nodes in the same cell but can only establish communication to one smart node at a time. Such a scenario is helpful, especially in emergency situations, where particular SUs may be restrained and attempt to outreach for any adjacent coverage and seek rescue. It is expected that CR can reach various unlicensed setups by using a predefined standard protocol dedicated for such situations.

7.2.2 Sensing Frame Structure

The FCC catered for the spectrum estimation intervals assignment at the start of each data frame of SUs in the IEEE 802.22 standard [230]. Consider a local spectrum estimation scenario comprising a number of PUs of unknown statuses and SUs working in the same bands of interest. A diagram showing the PUs statuses and SUs periodic sensing frame structures is given in Fig. 7.3, which also represents a spectrum handoff model [228].

The activity statuses of PUs are assumed unaltered during each complete frame of SUs. The sensing intervals in the front of SUs' frames are aimed at monitoring the

presence of PUs. The SUs are then to suitably decide on whether to transmit data or keep silent in the remaining portion of each frame. The sensing time is usually optimised in terms of achievable throughput determined by each technology. Throughputs are normally governed by the type of sensing mode, whether active (proactive) or quiet (silent or reactive) [90, 94, 95], threshold level and number of samples dedicated for sensing duration besides other requirements [231]. Sophisticated gears, such as MTM, can be used for the fine part of spectrum estimation, while the ED can be used for the coarse part in each frame [228].

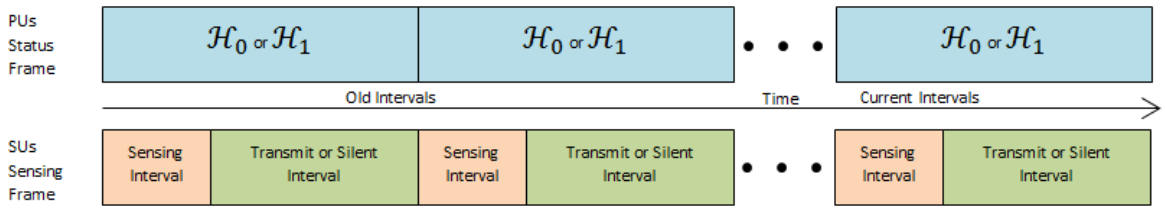


Fig. 7.3 Pictorial of the unknown PUs activity and the SUs periodic sensing frames structure.

The sensing period is selected to be 30 seconds in the IEEE 802.22 standard [8, 229]. Whereas the sensing frequency, channel detection and mobility times in addition to some other timing related parameters are also defined in the standard [229].

7.2.3 System Architecture

Fig. 7.4 portrays a generic architecture of SS in a wireless CRs that is used in this work. Suppose PUs and SUs are equipped with N_t and N_r transmit and receive antennas, respectively. Most building blocks of a CR system are typical in any wireless network, whether configured as a standalone or cooperative group. The SDR manifests the physical layers in the standard hierarchy of conventional operating systems and has controllability over all hardware components. It can interrogate the useful information from processed incoming sequences, gain knowledge and alter the hardware parameters to cope with the necessary surrounding changes. Also, the same CR structure can be used either for transmission or reception. Such a CR system is able to sweep all channels, or subbands, in the useful operating frequency band and store their status in its memory. As described above, the sweeping cycle has to be repetitive in order to

keep monitoring the status of PUs in each channel and thereby SUs can transit to any free channel consequently.

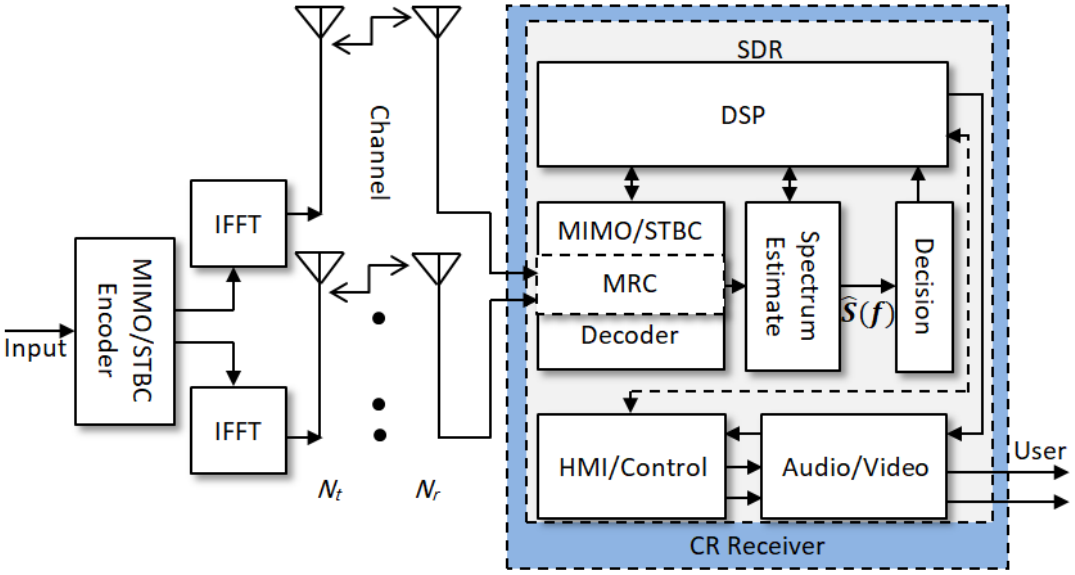


Fig. 7.4 Generalised architecture of SS in a wireless CRs.

It can be seen from the above diagram that the SS algorithm involves two main processing stages; the first is spectrum estimation to evaluate the spectrum content of the incoming signal, and the second is decision making in faith of a sufficient test statistic derived from the spectrum estimator in the first stage. An appropriately selected spectrum estimator is a key to the reliable detection of PUs' activities in the bands of interest. While on the other hand, the decision making policy is almost unified regardless of the nature of spectrum estimation being carried out. Therefore, care should be exercised while dealing with this application oriented task.

7.3 Problem Statement

SS is basically meant to distinguish between two cases; either the PU signal is active (present) or the PU signal is passive (absent), with minimum errors as possible. Denoting the hypotheses of the PU absence and the PU presence by $\mathcal{H}_0$ (noise only hypothesis) and $\mathcal{H}_1$ (signal plus noise hypothesis) respectively, this BHT problem for SISO settings can be expressed as below [2, 55, 90-93, 95-100, 101, 103-105, 142, 201]

$$\begin{aligned} \mathcal{H}_0: &\quad y(k) = n(k) \quad , \quad k = 1, \dots, K \\ \mathcal{H}_1: &\quad y(k) = hx(k) + n(k) \quad , \quad k = 1, \dots, K \end{aligned} \tag{7.1}$$

where $n(k)$ is the channel noise, h represents the channel coefficient which is assumed flat or block-fading, $x(k)$ is the PU signal, $y(k)$ is the signal observed by the SU receiver, and K is the sequence length. It is assumed that channel noise $n(k)$ is AWGN with zero mean and variance σ_n^2, i.e., $\mathcal{CN} \sim (0, \sigma_n^2)$. This model manifests an indirect, local and non-cooperative SS environment.

The task of SS is to decide whether the observation vector y was generated under $\mathcal{H}_0$ or $\mathcal{H}_1$. This can be achieved by forming a sufficient test statistic $\Lambda(y)$ based on the statistical characteristic of the received signal and comparing it with a prefixed threshold value η. This entails applying the classical BHT to decide on such spectrum occupancy. The BHT will have $\mathcal{H}_0$ to signify the idle state of PU and $\mathcal{H}_1$ signifies the active state of the PU. A sufficient statistical test Λ is essential in this case to conclude upon two separable decisions. For a predetermined threshold value η, the general test decision is conducted as per the following way, where the argument (y) has been omitted for simplicity

$$\begin{aligned} &\text{Decide: } \mathcal{H}_0 \quad \text{if } \Lambda \leq \eta \\ &\text{Decide: } \mathcal{H}_1 \quad \text{if } \Lambda > \eta \end{aligned} \tag{7.2}$$

Two probabilities are of importance; 1) the probability of detection P_d, which is the probability that the PU channel is correctly detected by the SU, and 2) the probability of false alarm P_{fa}, which represents the probability of a false detection of the PU when it is in the idle state, i.e., no transmission on the PU particular channel [2, 55, 90-93, 95-100, 101, 103-105]. The relationship between these two probabilities govern the performance metric of any SS design, and such relationship is commonly known as the receiver operation characteristic (ROC) [90, 92, 95, 96, 97, 98, 100, 102, 142, 201]. This is in fact a NP approach given that the process in hand has a well-defined statistical characteristic, and in this case the Gaussian distribution is taken for granted under such circumstances. Given an arbitrary sufficient test Λ based on the received signal features, and in the framework of NP optimal criterion, then these two probabilities are defined as below

$$P_{fa} = \Pr\{\Lambda > \eta | \mathcal{H}_0\}$$
$$P_d = \Pr\{\Lambda > \eta | \mathcal{H}_1\} \tag{7.3}$$

where η is a threshold calculated against the required constant false alarm rate (CFAR). Using the CFAR technique is a common practice in setting the threshold of spectrum estimators when the noise is of unknown power [92, 95, 98, 142, 201].

The design of decision, or detection, algorithms can be generally put under two main statistical frameworks; the classical and the Bayesian [92, 96, 142, 201]. In the classical, which is also known as deterministic, framework, either $\mathcal{H}_0$ or $\mathcal{H}_1$ is deterministically true, and the objective is to choose Λ and η so as to maximise P_d subject to a constraint on P_{fa} should be below a certain value; this is what the NP criterion is all about. While in the Bayesian framework, it is assumed that the detector randomly selects the true hypothesis in accordance to some a priori probabilities $\Pr\{\mathcal{H}_0\}$ and $\Pr\{\mathcal{H}_1\}$, and the objective hence is to minimise the Bayesian cost function. Despite the substantial difference between the two detection philosophies, both end up in a test of the forms (7.2) and (7.3). In the same manner of AMC, the Bayesian approach takes the typical LRT form below

$$\Lambda(\mathbf{y}) = \frac{p(\boldsymbol{y}, \mathcal{H}_1)}{p(\boldsymbol{y}, \mathcal{H}_0)} \tag{7.4}$$

Applying the LRT is challenging since the exact distribution given in (7.4) needs to be known. Obviously, the PDF of $\boldsymbol{y}$ under $\mathcal{H}_1$ is related to the distributions of the received signal, the wireless channel, and the noise, while the distribution of $\boldsymbol{y}$ under $\mathcal{H}_0$ is merely related to the noise distribution [96]. In order to apply the LRT, the knowledge of the channel as well as the signal and noise distributions should be provided in advance, which is difficult to grasp in reality. Several studies on the LRT approach for SS are accessible [232-234], however, owning to its simplicity and by embarking on the assumption that most channels undergo Gaussian statistics, only the NP approach is attempted in this work.

In light of the above, it turns out that the choice of sufficient test statistics and amenable decision attributions play a critical role in the SS design, which will be explored further in what follows.

7.4 Non-Parametric Spectrum Estimation

The approach of non-parametric spectrum estimation does not require a postulated model for the stochastic process under investigation, but assumptions regarding the unknown parameters are expected to be either known or given beforehand. Therefore, it can be considered as a blind spectrum estimation technique. The construction of such an estimator is mostly hinged on the FFT algorithms and, hence, is well perceived as computationally efficient. However, the modest performance, in some of the early estimation techniques like periodogram, in terms of consistency and biasing, can be enhanced by using several windowing or truncation methods. The spectrum estimation is applied on the incoming baseband signal to generate the PSD function, which signifies a sufficient test statistic and honoured to perform the BHT detection in this work.

7.4.1 Periodogram

The periodogram is a classical spectrum estimation technique that is readily implemented using the FFT. The estimate of PSD is achieved by forming an average of the discrete transforms of a windowed data sequence. The choice of window function is a compromise between an increased main-lobe width of the PSD estimate and the amount of side-lobe suppression that is obtained.

Let $y_w(k) = y(k)w_R(k)$ denote a windowed segment of samples obtained from the random process $y(k)$ at the receiver input, where the rectangular window function $w_R(k)$ contains K non-zero samples. The periodogram spectrum estimator (PSE) is then defined as the squared-magnitude FFT divided by K as follows [142, 201, 235]

$$\hat{S}(k) = \frac{1}{K}|\mathrm{FFT}(y_w)|^2 = \frac{1}{K}\left|\sum_{k=0}^{K-1} y_w e^{-j\omega k}\right|^2 \tag{7.5}$$

where the number of samples in the time and frequency domains must be normally equal to each other, which is K in this case. This is the PSD estimate and can be easily computed by applying the FFT on the squared magnitude of truncated signal as follows

$$y_w(k) \xrightarrow{\text{FFT}} Y_w(k) \rightarrow \frac{1}{N}|Y(k)|^2 = \hat{S}\left(e^{j2\pi k/K}\right) \tag{7.6}$$

7.4.2 Filter Banks

The concept of FBks application to spectrum estimation is not new. The widely used periodograms and the related algorithms, such as weighted overlapped segment averaging (WOSA), may be viewed as FBk spectrum estimators (FBkSEs) with relatively simple prototype filters [121, 236]. Let $h_l(k)$ be a FIR filter of length K that is defined as follows [235]

$$h_l(k) = \frac{1}{K} e^{j\omega_l k} w_R(k) = \begin{cases} \frac{1}{K} e^{j\omega_l k} & , \quad 0 \le k < K \\ 0 & , \quad \text{otherwise} \end{cases} \tag{7.7}$$

This is a bandpass filter (BPF) with $(\sin x/x)$ response function centred around ω_l and has a small enough bandwidth of $\Delta\omega = 2\pi/K$. If the incoming signal $y(k)$, considered a WSS random process and filtered with $h_l(k)$, then the output process is given below

$$\tilde{y}_l(k) = y(k) * h_l(k) = \sum_{i=k-K+1}^{k} y(k) h_l(k-i) = \frac{1}{K} \sum_{i=k-K+1}^{k} y(k)\, e^{j(k-i)\omega_l} \tag{7.8}$$

After certain approximations, the PSD estimate at the l^{th} subband of a FBk is obtained as

$$\hat{S}_l(k) = \frac{1}{K} \sum_{k=0}^{K-1} [|\tilde{y}_l(k)|^2] = \frac{1}{K} \left| \sum_{k=0}^{K-1} y(k) e^{-j\omega_l k} \right|^2 \tag{7.9}$$

and the total PSD is given by averaging the output PSD of all L filters in the FB scheme

$$\hat{S}(k) = \frac{1}{L} \sum_{l=1}^{L} \hat{S}_l(k) \tag{7.10}$$

Thus, the periodogram may be viewed as the estimate of the power spectrum that is formed using a bank of BPFs as shown in Fig. 7.5. The periodogram has such a FBk feature "built into it" and hence it is not necessary to build one.

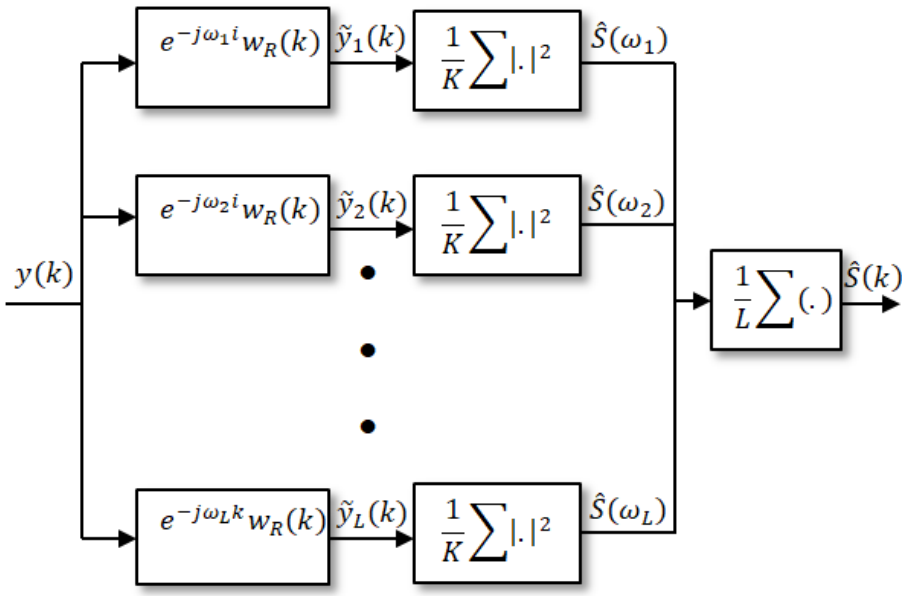


Fig. 7.5 FBk interpretation of the periodogram.

A different treatment of the above concept can also be worked out. The assumption of having the prototype filter of an N band FBk is given by [236]

$$H(z) = \sum_{k=0}^{M-1} h(k)z^{-k} \tag{7.11}$$

Assuming that the i^{th} band of the FBk has the centre frequency $\frac{2\pi i}{N}$, its transfer function is

$$H_i(z) = \sum_{k=0}^{M-1} h(k)W_N^{-ik}\, z^{-k} \tag{7.12}$$

where $W_N = e^{\frac{-j2\pi}{N}}$. Let $M = KN$, the above equation may then be rearranged as follows

$$H_i(z) = \sum_{l=0}^{N-1} z^{-l}E_l(z^N)W_N^{-il} \tag{7.13}$$

where the l^{th} polyphase component of $H(z)$ is given by

$$E_l = h(l) + h(l+N)z^{-1} + \cdots + h(l+(K-1)N)z^{-K} \tag{7.14}$$

The polyphase structure resulting from the direct application of (7.12) is depicted in Fig. 7.6 below. If N is appropriately selected as a composite number, the FBk can then

be efficiently realised using the IFFT algorithm. This is nothing but a moving window spectrum analyser implemented over the entire K-length sequence, updated every N-point FFT samples to assimilate the number of N bands being used in the FBk. The resolution of the spectrum analyser can be improved by increasing the number of N bands, but that of course will be at the expense of increasing the computational load.

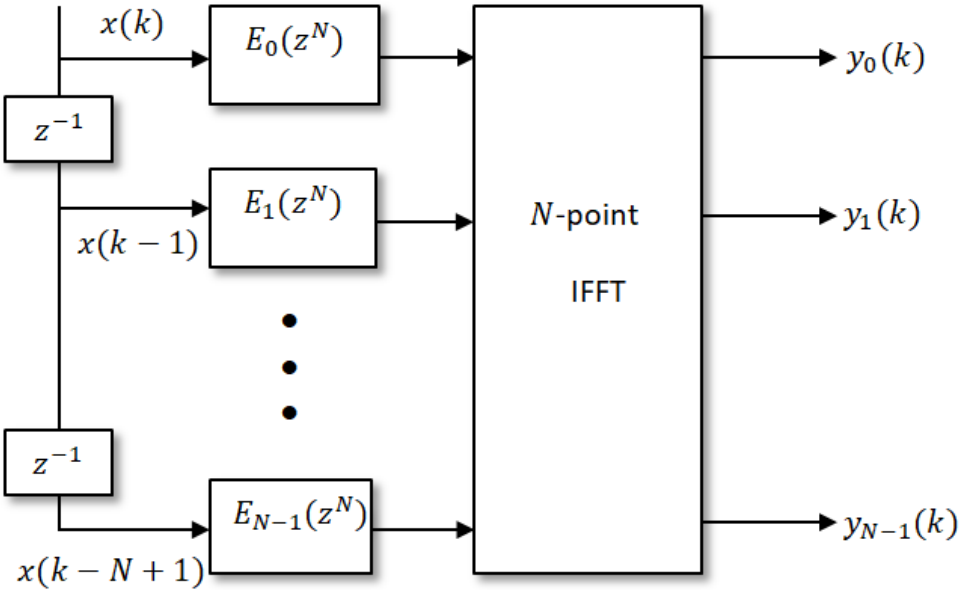


Fig. 7.6 Polyphase realisation of an N-band FBk.

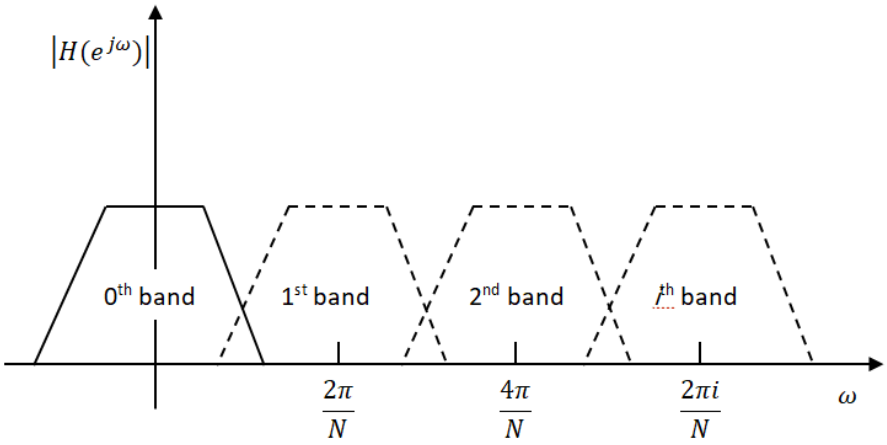


Fig. 7.7 Graphical presentation of FBk.

Given the above, the implementation of the FBkSE has now become more obvious. The input signal is passed through a bank of filters and the output power of each filter is measured as an estimate of the power spectrum over the associated subband. The final

PSD is the average of all powers over all bands. The realisation of FBks is often based on prototype filters. The prototype filter is a LPF that is also used to realise the 0^{th} band of the FBk. Other bands are realised through a frequency shift of the prototype filter. A graphical presentation of this concept is shown in Fig. 7.7 above.

7.4.3 Multitaper Method

In the approach of MTM for spectral analysis, the data record is not segmented as in WOSA; rather, several different tapers are chosen usually from the family of sequences known as DPSS. These tapers are called Slepian tapers (sequences), and named after the first researcher who elaborated on their effects [1, 106, 236, 237]. Each taper is applied, in turn, to the full data record and a periodogram computed via the FFT. Finally, the periodograms are averaged to produce a spectrum estimate. The performance of this estimator depends on the data record length, the number of tapers, and the tapers themselves. Since each single periodogram is computed over the entire data stream, the MTM can exhibit greatly improved performance in contrast to WOSA [106, 237].

The MTM algorithm for spectrum estimation was claimed to be an approximation to MLE for PSD function [90]. The MTM for spectrum estimation can also be thought as a FBkSE with multiple FBks [91, 92]. All the advantageous features that MTM enjoys are primarily attributed to the Slepian tapers because they enjoy two important features. First, their FFTs have the maximal power concentrated in the main lobes while the spectrum leakage in the undesirable OOBs is kept to the least extent when applied for spectral estimation. Second, they are orthogonal to each other, and that translates into the estimate outputs employing different tapers are uncorrelated providing that the spectrum variation over each subband is insignificant. Averaging these individual estimates will therefore produce a MVUB estimate. Due to the two characteristics of the Slepian sequences, the MTM for spectrum estimation (MTSE) has been shown to be nearly optimal in the sense that it almost achieves the CRLB for a nonparametric spectral estimator [90, 93, 95]. Hence, the MTSE has been highly voiced by Haykin as a promising power spectrum estimation for wideband SS [1, 106, 140, 141, 237]. Recent studies have echoed such reference and promoted the robust MTM for SS in CR applications [92, 93, 237].

The MTM linearly expands the part of the time series, in a fixed bandwidth extending from $f - B$ to $f + B$ (centred on some frequency f), in specialised Slepian tapers. The remarkable property of these tapers is that their Fourier transforms have the maximal energy focused in the bandwidth $2B$ under a finite sample-size constraint. This property, in turn, permits trading spectral resolution for improved spectral characteristics, that is, reduced variance of the spectral estimate without compromising the bias of the estimate.

The steps below are for the MTSE application as given in [1, 106, 237] and partially in [92]

- An orthonormal sequence of Slepian tapers $\{v_k^l\}_{k=0}^{K-1}$.
- A corresponding FFT set with $l = 1, \dots, L$ and L is the total number of tapers

$$Y_l(k) = \frac{1}{K}\sum_{k=0}^{K-1} y(k) v_k^l e^{-j\omega_l k} \tag{7.15}$$

- A normalised spectrum estimate, based on the first few eigenspectra that exhibit the least sidelobe leakage, is

$$\hat{S}(k) = \frac{\sum_{l=1}^{L} \psi_l |Y_l(k)|^2}{\sum_{l=1}^{L} \psi_1} = \sum_{l=1}^{L} \alpha_l \, |Y_l(k)|^2 \tag{7.16}$$

where ψ_l is the eigenvalue of the l^{th} eigenspectra, $L \leq 2KB$ bounds the number of tapers and defines the DoF for estimating the variance control. The weighting factors

$$\alpha_l = \psi_l / (\psi_0 + \cdots + \psi_L) \tag{7.17}$$

are used for the eigenspectra significance quantification. Using the KLT to maximise the signal power, the following eigen equation can be used to determine the Slepian tapers

$$\boldsymbol{R}\boldsymbol{v_l} = \psi_l \boldsymbol{v_l} \tag{7.18}$$

where $\boldsymbol{R}$ is the $K \times K$ positive-definite Toeplitz autocorrelation matrix and the $(i,j)^{\text{th}}$ entry of this kernel matrix is defined below

$$\boldsymbol{R}_{i,j} = \frac{sin(2\pi B(i-j))}{\pi(i-j)} \quad , \quad i,j = 1,2,\dots,K \tag{7.19}$$

The eigenvalues ψ_l range between 0 and 1, and are organised in descending order such that $\psi_0 \geq \psi_1 \geq \cdots \geq \psi_L$. The first $L \approx 2KB$ of them are dominated, close to 1, whereas the rest are negligible. Moreover, the tapers of lower order have capability for much stronger energy concentration than their higher order counterparts, suggesting that it suffices to use the first L tapers for spectral estimation. The choice of the parameter time-bandwidth product $L_0 = KB$ is a tradeoff between spectral resolution and variance, where K is the time sample size equivalent to the FFT window, and B is the bandwidth normalised by the sample rate. Typically, $3 \leq L_0 \leq 6$ are dominant data tapers out of the extensive range of numbers to perform the estimation. The block diagram of this MTSE is shown in Fig. 7.8.

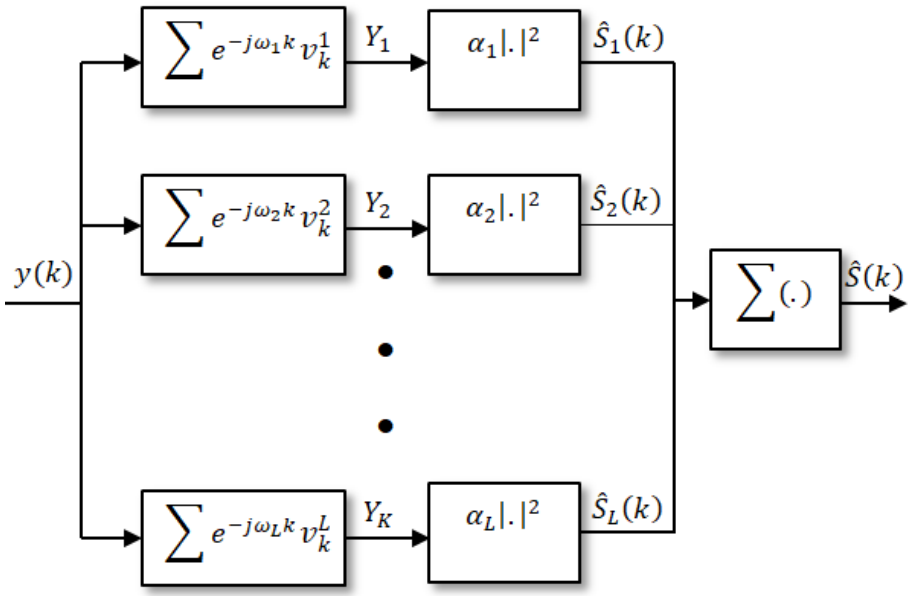


Fig. 7.8 MTSE block diagram.

The trend of various Slepian tapers is illustrated in Fig. 7.9 using 6 tapers and L_0 equals 0.0625 and 1, respectively. In contrast to the first value, the second value is distinctive since it indicates the margin between low and rapid amplitude fluctuations of such tapers. This will accordingly constitute more tapers to be accounted for their energy congregates.

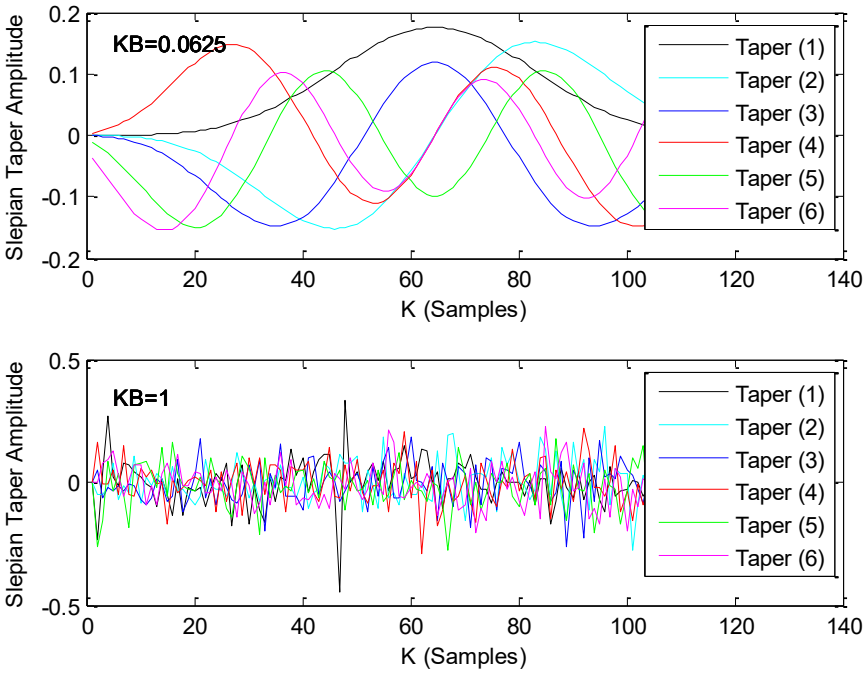


Fig. 7.9 Representation of 6 Slepian tapers with two different KB values.

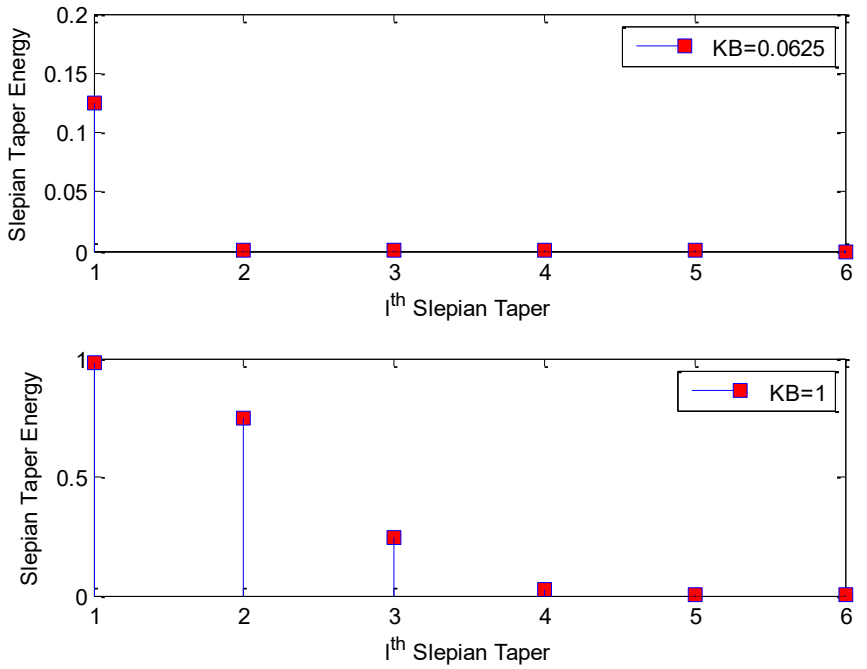


Fig. 7.10 Energy concentration of 6 Slepian tapers with two different KB values.

It is evident from Fig. 7.10 above that increasing the number of tapers would not be rewarding for the MTSE algorithm because most of the effective energy is focused in the low-order tapers. These lower-order tapers impose greater control on the spectrum bias, and the opposite is true for the higher-order tapers. It is obvious that the behaviour of Slepian tapers act as a BPF, as given by the averaging effect in (7.15) and (7.16), and hence deliver sharp spectrum truncation and better smoothing. Increasing the number of

tapers will have energy dispersed over many orthogonal elements. As the half time-bandwidth increases from 0.0625 to 1, the tapers eventually develop more erratic behaviour as lots of energy will be shifting to the higher-order tapers.

7.4.4 Quadratic Form Representation

The sufficient test statistic resulting from applying the MTSE is the PSD estimate. By combining (7.15) and (7.16) yields the following expression for such estimate

$$\hat{S}_y(f) = \frac{1}{LK}\sum_{l=1}^{L}\sum_{k=0}^{K-1}\alpha_l\left|y(k)v_l(k)e^{-j2\pi fk}\right|^2 \tag{7.20}$$

where f is the frequency argument and can be taken as k for discrete sequence $\forall k \in [1, K]$. Tackling this expression is not straightforwardly conceivable in the sense that there is no ready-made statistical model that can be easily fitted to it. Therefore, other alternative reckonings need to be sought in order to further simplify the problem at hand as possible.

The quadrature form representation suggests itself as one of the best candidates to render the expression of (7.20) furtherly simplified and land a discernible statistical metric accordingly. It often occurs that the PDF of $y = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A}$ is a positive definite or semidefinite symmetrical Hermitian $K \times K$ matrix and $\boldsymbol{x}$ is an $K \times 1$ Gaussian vector, is virtually required [142, 201, 235]. Tailoring such quadrature form to suit (7.20) greatly simplifies the analysis and produces tractable statics necessary for the BHT detection stage.

Acknowledging the above need, several studies endeavoured to facilitate the quadrature simplification of the MTSE under various work scenarios [238-241]. To have traceable illustrations, a new set of vectors is defined to exploit the quadratic matrix formation. Extending the approach of [238-241], the vector variables below are assumed

$$\begin{aligned} \boldsymbol{v}_l &= [v_l(1), v_l(2), \dots, v_l(K)]^{\mathrm{T}} \\ \boldsymbol{y} &= [y(1), y(2), \dots, y(K)]^{\mathrm{T}} \\ \mathbf{a} &= \left[e^{j2\pi f}, e^{j2\pi f2}, \dots, e^{j2\pi fK}\right]^{\mathrm{T}} \end{aligned} \tag{7.21}$$

where $\mathbf{b}_l = \boldsymbol{\upsilon}_l \odot \mathbf{a}$ and $\odot$ refers to the Hadamard product. Substituting (7.21) in (7.15) reveals $\boldsymbol{Y}_l(f) = \mathbf{b}_l^{\mathrm{H}}\boldsymbol{y}$, and when used together with the Slepian tapers alters (7.20) to have the following PSD estimate using quadrature form approximation

$$\hat{S}(f) = \boldsymbol{y}^{\mathrm{H}}\boldsymbol{\Omega}\boldsymbol{y} \tag{7.22}$$

where $\boldsymbol{\Omega}$ is $L \times L$ idempotent matrix defined as $\boldsymbol{\Omega} = \sum_{l=0}^{L-1} \alpha_l \mathbf{b}_l \mathbf{b}_l^{\mathrm{H}}$ and $(.)^{\mathrm{H}}$ is the Hermitian operator. The above scalar PSD expression can be viewed as the sum of individual PSDs calculated over dominant tapers. It is notable that the IFFT order is deliberately assumed equivalent to the length of data sequence for ease of treatment and, importantly, to avoid adverse computations in the assessment settings by simulation.

The multi-space projection of the above geometrical vectors has been demonstrated in [242]. It also has indicated that the MTSE uses a sliding window (called bin) of certain resolution to sweep the received signal. Such a sliding window is supported by the frequency-shifted version of $\boldsymbol{\upsilon}_l$, which takes the shape $\mathbf{b}_l$ in the quadrature form, and as the baseband $\boldsymbol{\upsilon}_l$ has a resolution of B, its frequency shifted version $\mathbf{b}_l$ has a double resolution bin width. As the bin moves to scan the received vector $\boldsymbol{x}$, the received vector $\boldsymbol{x}$ is projected onto the corresponding bin subspace. The wider the frequency resolution bin, the more the dimensions are available, and hence the smaller the bias is achieved. This helps the MTSE improve the estimation resolution accordingly [242].

7.5 Decision Statistics

The statistical expressions are formulated for the hypothesis test, threshold and sensing frame settings that constitute the core measure metrics for the SS performance.

7.5.1 Hypothesis Test

Since the PSD estimate $\hat{S}(f)$ denotes the sufficient test Λ, and in the framework of NP optimal norm, the pair of probability metrics (7.2) and (7.3) can thus be deduced as follows

$$P_{fa} = \Pr\{\hat{S}(f) > \eta | \mathcal{H}_0\}$$
$$P_d = \Pr\{\hat{S}(f) > \eta | \mathcal{H}_1\} \tag{7.23}$$

To examine the distribution of the random vector $\hat{S}(f)$ under both hypotheses, it is useful to invoke a new term called the partial test statistics [234, 244]. The distribution of quadrature form representations is well known to be of Chi-Square nature [142, 201, 235]. Featuring the approach of partial test statistics is fairly hosted here to unfold the impact of various transmission settings, such as MIMO, STBC and OFDM. It offers a thoughtful exposition on how the various parameters are statistically accrued.

The expression of partial test statistics was firstly exploited for ED algorithms in some recent studies [243, 244]. One possible rational is by evoking $\mathcal{R}$ separate channel paths and each is loaded with K samples of *iid* complex Gaussian noise with zero mean and σ_n^2 variance. Then each taper branch will convert each sample into central Chi-square variable χ^2 and summing over L tapers results in χ_{KL}^2 of KL DoF, which is assigned for the $\hat{S}(f)$ statistics. In the case of MIMO and STBC settings featuring SVD, the summation is conducted over all the decomposed channel paths, and the final variable will hence be of $\chi_{\mathcal{R}LK}^2$ statistics. Then, and in similar approach to [244], but approximating the sum of all channel eigenvalues to 1, the above BHT statistics for the MTSE-MIMO are compiled as below

$$\text{Decide: } \mathcal{H}_0 \text{ if } \hat{S}(f) \approx^{\mathrm{d}} \sigma_n^2 \chi_{\mathcal{R}LK}^2$$
$$\text{Decide: } \mathcal{H}_1 \text{ if } \hat{S}(f) \approx^{\mathrm{d}} (\sigma_s^2 + \sigma_n^2) \chi_{\mathcal{R}LK}^2 \tag{7.24}$$

where $\approx^{\mathrm{d}}$ stands for "equal in distribution". The probabilities can hence be further simplified by assuming a large number of combined random sequences and channels. Given this large number theory, or equivalently the CLT, the probability approximations given in [245-250] can be modified as follows

$$P_{fa} = \mathrm{Q}\left\{\frac{\eta - \mathcal{R}LK\sigma_n^2}{2\sigma_n^2\sqrt{\mathcal{R}LK/2}}\right\}$$
$$P_d = \mathrm{Q}\left\{\frac{\eta - \mathcal{R}LK(\sigma_s^2 + \sigma_n^2)}{2(\sigma_s^2 + \sigma_n^2)\sqrt{\mathcal{R}LK/2}}\right\} \tag{7.25}$$

and after straightforward manipulation to involve the SNR perspective converts the probabilities pair of (7.25) into the following forms

$$P_{fa} = \mathrm{Q}\left\{\frac{\acute{\eta} - \mathcal{R}LK}{2\sqrt{\mathcal{R}LK/2}}\right\}$$
$$P_d = \mathrm{Q}\left\{\frac{\acute{\eta} - \mathcal{R}LK(\gamma+1)}{2(\gamma_s+1)\sqrt{\mathcal{R}LK/2}}\right\} \tag{7.26}$$

where the revised threshold $\acute{\eta} = \eta/\sigma_n^2$ and the noise remains of a WSS type, $\gamma_s = \mathrm{SNR} = \sigma_s^2/\sigma_n^2$, and $\mathrm{Q}(.)$ is the right tail cumulative distribution function and in terms of the complementary error function $\mathrm{Q}(.) = \frac{1}{2}erfc(\frac{.}{\sqrt{2}})$ for any Gaussian random variable. The signal and noise variances are assumed known and the dependence of both probabilities on the threshold value $\acute{\eta}$ becomes evident. Furthermore, all constants are discarded and the channel and taper eigenvalues are approximated to 1 for simplicity.

It is now important to assess the different impacts of STBC scheme on the MTSE probability attributes given above. The overall received SNR value generated by each scheme is of critically direct influence on the MTSE general performance. Relating back to the discussion in Chapter Three, it is obvious that the total amount of SNR, which is contributed by the real and imaginary parts of STBC complex symbols, produces an effective SNR that is M times higher than that of the normal MIMO scheme. Moreover, not only the STBC decouples the transmission channels into independent parallel paths of scalar dominant eigenvalues, the STBC also transforms space-time channel into M parallel and independent AWGN channels associated with each part of the complex symbol [118, 119]. Putting the above in mathematical footings, and guided by [244-250] for the mean and variance of signal and noise components and the approximations of other statistical attributes, the STBC reshapes (7.26) to yield the following MTSE-STBC forms

$$P_{fa} = \mathrm{Q}\left\{\frac{\acute{\eta} - \mathcal{R}MLK}{2\sqrt{\mathcal{R}MLK/2}}\right\}$$
$$P_d = \mathrm{Q}\left\{\frac{\acute{\eta} - \mathcal{R}MLK(\gamma+1)}{2(\gamma_s+1)\sqrt{\mathcal{R}MLK/2}}\right\} \tag{7.27}$$

Recall from the discussion in Chapter Three, the STBC obviously has extra power advantage over an ordinary MIMO system due to the built-in coding process. Such additional gain is attributed to the direct reliance on the M-ary constellation symbols observing the same channel eigenvector at a time. Engaging the STBC hence doubles the coding gain compared to that of MIMO system irrespective of diversity gain and channel fading conditions. This conforms to the coding gain of around 3 dB is readily achievable for 2×2 STBC [55, 118], which increases by $10log_{10}M$ for higher order codes. Also, the number of eigenvalues for the Alamouti case is four times higher than that for the general MIMO system due to the STBC coding is performed over two time intervals, and the observation is divided into real and imaginary parts. In addition to that, the STBC practice is considered robust in case of poor channel reporting due to fading effects. A receiver hence not necessarily always needs the CSI to perform the spectrum estimation duties and can work in a blind or semi-blind fashion in case involving STBC in its structure [55, 118].

Attempting closed-form expressions for the SS probability metrics involving OFDM signalling can also be feasible [92, 232, 243, 246]. Among other various OFDM detectors, essential approximations of the GLRT and using the MLE to interrogate the CSI unknowns was considered. The MLE approach also proved useful in a blind SS based on the GLRT algorithm for OFDM detection [251]. It is worth to mention, the MLE algorithm was also facilitated for SS in MIMO settings to feed the GLRT with necessary estimates of the unknowns [233, 234]. Despite it being a well understood and mature technology, however, some other studies have pointed out that the OFDM suffers spectrum leakage due to large sidelobes, and hence the OOB generated energy can be significant [93, 98, 236]. Such large spectral sidelobes can be problematic for SS in CR systems. The analysis proposed in what follows generally harmonises with the same opinion of the OFDM bad influence on the SS performance. However, a rather novel route is presented here which makes the given analysis unique and different from other treatments elsewhere.

Suppose a link transmission is taking place using OFDM with N_c subcarriers arranged in a MIMO configuration over a flat or block fading channel. The complex baseband received sequence over $n_c = \{0, 1, \dots, N_c\}$ subcarrier is denoted by $\boldsymbol{Y}_{n_c} = \boldsymbol{H}_{n_c}x_{n_c} + \boldsymbol{N}_{n_c}$, where all variables are as defined before in the context, but now defined

over each subcarrier, or frequency bin, of the OFDM signalling. The same SVD procedure is also applicable in this situation and hence will be exploited in achieving the final metric results here. Consider now the effect of the entire frequency bins on the probabilities as given above for each individual bin. For such a situation, the problem can be reduced to the detection of an unknown phase or frequency signal [201], where the total P_{fa} and P_d easily lend themselves to have the following forms

$$\begin{aligned} P_{fa} &= N_c P_{fa}(bin) = N_c \mathrm{Q}\left\{\frac{\acute{\eta} - \mathcal{R}LK}{2\sqrt{\mathcal{R}LK/2}}\right\} \\ P_d &= \mathrm{Q}\left\{\left(\frac{1}{(\gamma_s+1)}\right)\mathrm{Q}^{-1}\left(\frac{P_{fa}}{N_c}\right) - \sqrt{\frac{\mathcal{R}LK}{2}}\gamma_s\right\} \end{aligned} \tag{7.28}$$

It is evident that P_{fa} increases about linearly with respect to the number of bins examined, which means further performance degradation as the OFDM order increases. That means, the more subcarriers, the more subbands and hence the more OOB spectrum leakage is expected. Such application is typically known as a wideband or multiband SS [252, 253], which is different than an individual band or subchannel SS associated to each bin or carrier frequency commonly encountered elsewhere in the literature.

7.5.2 Threshold Setting

The performance of MTSE for SS can be measured by the probability metrics P_{fa} and P_d. These metrics have high dependency on the threshold value $\acute{\eta}$ assuming that both the signal and the noise variances, σ_s^2 and σ_n^2, are either known or estimated. A common method for setting the threshold value is to be assigned against a prefixed CFAR, which is equal to P_{fa} [92, 95, 96, 239, 242, 246, 249, 250]. Under hypothesis $\mathcal{H}_0$, the threshold value for the MIMO and the STBC can be deliberated as given below, respectively

$$\acute{\eta} = \mathcal{R}LK\left(1 + \frac{2\mathrm{Q}^{-1}(P_{fa})}{\sqrt{\mathcal{R}LK/2}}\right) \tag{7.29}$$

$$\acute{\eta} = \mathcal{R}MLK\left(1 + \frac{2\mathrm{Q}^{-1}(P_{fa})}{\sqrt{\mathcal{R}MLK/2}}\right)$$

Another approach for setting the threshold value is by using the constant detection rate (CDR), or P_d, under hypothesis $\mathcal{H}_1$ yields the following expression for the MIMO and the STBC, respectively

$$\begin{aligned}\acute{\eta} &= \mathcal{R}LK(\gamma_s + 1)\left(1 + \frac{2\mathrm{Q}^{-1}(P_d)}{\sqrt{\mathcal{R}LK/2}}\right)\\ \acute{\eta} &= \mathcal{R}MLK(\gamma_s + 1)\left(1 + \frac{2\mathrm{Q}^{-1}(P_d)}{\sqrt{\mathcal{R}MLK/2}}\right)\end{aligned} \tag{7.30}$$

while working out the same for the MIMO-OFDM settings yields the following thresholds CFAR and CDR, respectively

$$\begin{aligned}\acute{\eta} &= \mathcal{R}LK\left(1 + \frac{2\mathrm{Q}^{-1}(P_{fa}/N_c)}{\sqrt{\mathcal{R}LK/2}}\right)\\ \acute{\eta} &= \mathcal{R}LK(\gamma_s + 1)\left(\gamma_s + \frac{1}{(\gamma_s + 1)} + \frac{2\mathrm{Q}^{-1}(P_d)}{\sqrt{\mathcal{R}LK/2}}\right)\end{aligned} \tag{7.31}$$

The decision of setting threshold value is highly pertinent to the CR system or CRN strategy. Values based on CDR are aimed to provide higher protection for PUs against intrusions from opportunistic SUs, and values based on CFAR are applied in favour of allowing SUs better access opportunities to the frequency spectrum. In the first scenario, the CFAR value shall be chosen to be relatively high, or CDR is low, while in the second the opposite is true. Therefore, some adaptation techniques could prove helpful in determining the adequate threshold values under dynamically changing environments and as per spectrum access policy in place. This work is concerned with fixing the threshold value against a preassigned CFAR value only, and the AT will be applied using one signalling scheme as the rest share the same concept exactly.

7.6 Adaptation Techniques

In the world of dynamically changing environment, having the SS parameters fixed to satisfy single targeted CFAR or CDR limits is not adequate for flexible CR

operations. Therefore, the requirement to adapt the various SS parameters to their environment and provide resilience under adverse conditions has lately become knocking at all levels.

The issue of AT and OFD is a challenging task for the seamless operation of SS in CR systems [199, 230, 245-250, 253]. The SS performance at any point in time is highly determined by the values of performance metric probabilities, which in turn are reliant on the threshold and frame duration settings. A suitably selected detection threshold and frame length can minimise SS error, furnish PUs with proper protection, and improve spectrum utilisation efficiency. Aiming at the targeted probability and threshold metrics, a subtle OFD necessary for SS in CRNs needs to be closely examined, as emphasised in earlier studies [231, 245, 255]. Furthermore, the 2 seconds evacuation time as per the IEEE 802.22 [8, 229, 235] places a heavy constraint on CR systems to have efficient designs aimed for maximising throughputs during the active and passive sensing modes [231, 253-256].

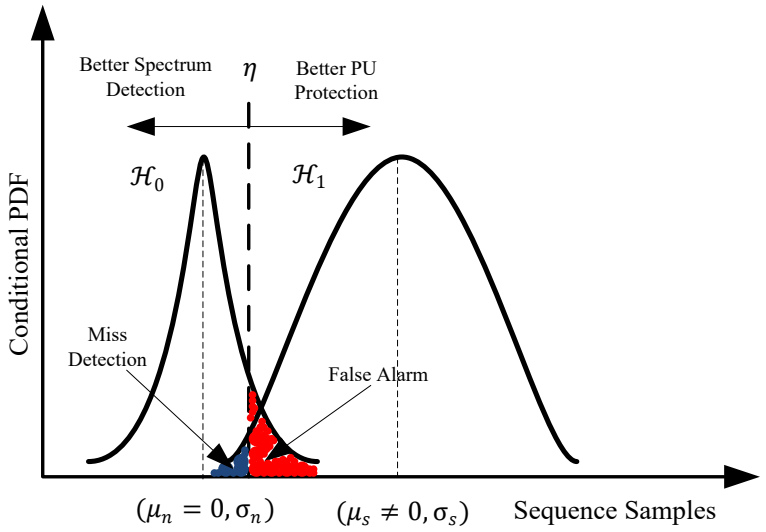
The threshold function is a typical convex optimisation problem explicitly expressed in two variables, namely; the missed detection P_{md} and false alarm P_{fa} probabilities [249, 250]. The SS performance at any point in time is highly determined by the values of these probabilities, which in turn are dependent upon the threshold value. A perfect knowledge of the noise power is essential to administer the SS decision threshold. Failure in knowing the noise will not result in the desired P_{fa} and renders the SS performance susceptible to errors, accordingly. A suitably selected decision threshold can minimise SS error, furnish PUs with proper protection, and improve spectrum utilisation efficiency. The higher threshold value, the better P_d and protection against indivertible interference to PUs is achieved. While the less threshold value, the higher P_{fa} and less PUs protection is generated. Therefore, an adequate threshold adaptation policy is mandatory for tradeoff between the two situations. On the other hand, the OFD has an inverse relationship with respect to the different SS parameters. Both, however, the AT and OFD are largely reliant on the statistical signalling features. This work is pioneered to present the analysis of these two critical metrics using the MTSE for SS for the first time, where the treatment of which is applicable to any kind of signalling schemes without any major alterations.

7.6.1 Threshold Criterion

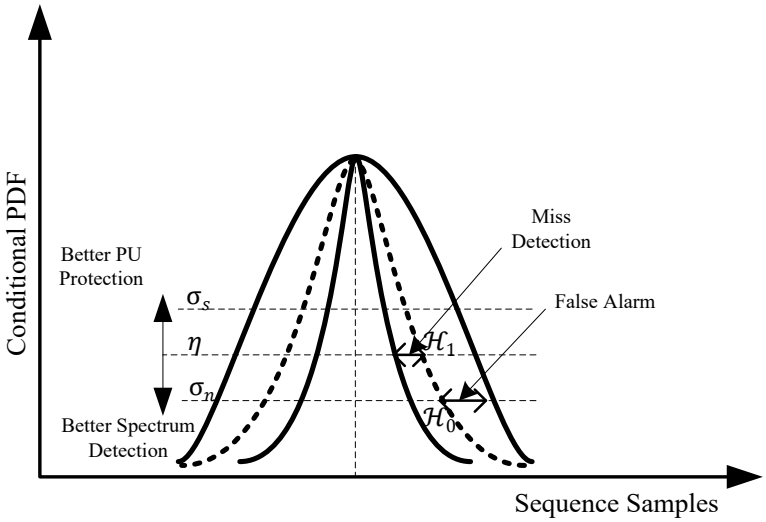
Before delving into the AT and OFD analysis, let the signalling statistics be examined first. Two particular situations are worthy to distinguish as can be observed in Fig. 7.9. Although a zero mean noise is normally valid in both situations, however, the signal sequence may tell a different story. This is largely pertinent to the signalling scheme being used for transmission, which can be either of zero mean or not irrespective of variance (or power). Fig. 7.11.a provides a statistical behaviour commonly reflected by many studies in the literature. Despite the signal appearing to have a non-zero mean value, however, an elaboration on such stance is almost dismissed in many studies, such as in [13]. It is notoriously known that the realism of such a situation is of practical limitations, as signalling of non-zero means does not meet certain transmission and processing design requirements. Obviously, the signal and noise variances could be different.

Furthermore, Fig. 7.11.a also portrays the false alarm P_{fa} and miss detection $P_{md} = 1 - P_d$ probabilities and how they can be controlled by AT. The typical purpose of most SS techniques is to reduce these two probabilities, as best as possible. However, this is not an easy task to achieve as these two probabilities behave freely from each other. Stretching the threshold value up or down will have a tandem effect on the PU protection and efficient spectrum utilisation, whether favourably or not. A tradeoff is essential in such challenging conditions, subject to the SS design and performance criteria.

On the other hand, Fig. 7.11.b seems more realistic. It shows zero mean values for both the signal and noise, which mostly resembles practical conditions. They differ only in variances. Imagine a statistical distribution is trended to emulate threshold, which lies between the other two that belong to signal and noise as shown in the same figure. The interpretation of increasing or decreasing threshold on the probabilities' metrics applies exactly in the same manner described above. It is imperative to assert that the analysis and derivations given in this work applies equally to both parts of Fig. 7.11. The SNR can further unravel this issue as given below.



(a)



(b)

Fig. 7.11 Signalling statistics (a) non-zero mean (b) zero mean.

The SNR value has a radical influence on P_d, or alternatively P_{md}, while the P_{fa} can be changed by the noise variance only and hence can be fixed to any desired value independently from P_{md}. As it was shown in Chapter Three, the MIMO, as well as orthogonal STBC, decouples the spatiotemporal channel into parallel and independent scalar channels, each with SNR expressed as given in (3.6). Rearranging and ignoring the channel envelop, the SNR may be reduced to have the shape $\gamma_s = (\mu_{s+}^2 \sigma_s^2)/\sigma_n^2$.

The above clearly shows the link between the first two statistical moments and the MIMO and STBC signal power design, as was generalised in Fig. 7.11 a & b, respectively. Therefore, articulation on the actual statistical behaviour of the signal and noise attributes and their intervention to analysis can be proved instructive and cannot be merely undeclared. Having said that, it is important to reiterate again that the MTSE-STBC discussion offered in this work is still applicable, whether the signal has an average value of zero or not.

7.6.2 Dynamic Adaptive Threshold

Generally, classical SS designs are based on either the CFAR or the CDR performance measures, to control P_{fa} and to control P_d, respectively. The CFAR maintains better spectrum reuse, while the CDR preserves better protection for PUs against interferences from undesired intruders. Both methods, however, produce constant performance indicators regardless of the SNR or noise power variations. In order to make it more rewarding for both PUs and SUs, a threshold adjustment methodology needs to be conveniently devised to achieve the best tradeoff between P_d and P_{fa} under various operational conditions.

Several studies suggested to use the weighted tradeoff principle to minimise the decision error probability of SS [199, 245-250], which is also adopted here. The optimum threshold value has to be chosen to minimise the following convex function of total error decision probability

$$P_E = \alpha P_{fa} + (1-\alpha)P_{md} \tag{7.32}$$

where a tradeoff formed between P_{fa} relative to that of P_{md} can be attained with respect to a predefined weighting, or spectrum utilisation, factor, within the range $0 < \alpha < 1$. Substituting (7.32) in (7.26) for MIMO settings yields

$$P_E = \alpha \mathrm{Q}\left\{\frac{\acute{\eta} - \mathcal{R}LK}{2\sqrt{\mathcal{R}LK/2}}\right\} + (1-\alpha)\left(1 - \mathrm{Q}\left\{\frac{\acute{\eta} - \mathcal{R}LK(\gamma_s+1)}{2(\gamma_s+1)\sqrt{\mathcal{R}LK/2}}\right\}\right) \tag{7.33}$$

This solution of this convex optimisation problem can be attended by using different methods. Examples may include the gradient method using Lagrange multipliers [199, 245], or dispersed chirp-z transform [247]. The most common, however is by taking the

first derivative of (7.33) with respect to $\acute{\eta}$ and equating it to zero subject to the second derivative being greater than zero [245, 246, 248-250]. After substituting $\mathrm{Q(z)} = \frac{1}{2\pi}\int_z^{\infty} e^{-\frac{t^2}{2}}\,dt$ and rearrangement, this reduces to the following quadrature equation

$$\frac{(\gamma_s+2)}{2(\gamma_s+1)}\acute{\eta}^2 - \sigma_n^2\acute{\eta} - \frac{2\sigma_n^2(\gamma_s+1)}{\gamma_s\mathcal{R}LK} ln\left\{\frac{\alpha}{(1-\alpha)}(\gamma_s+1)\right\} = 0 \tag{7.34}$$

and solving for the positive threshold gets as below

$$\acute{\eta} = \frac{1+\sqrt{1+\frac{4}{\mathcal{R}LK}\left(1+\frac{2}{\gamma_s}\right) ln\left(\frac{\alpha}{(1-\alpha)}(\gamma_s+1)\right)}}{(\gamma_s+2)/(\gamma_s+1)} \tag{7.35}$$

and for STBC $\mathcal{R}LK \rightarrow \mathcal{R}MLK$. As $\mathcal{R}LK$ tends to be a very large value approaching infinity, this was suggested to be approximated as [249]

$$\acute{\eta} \approx 2(\gamma_s+1)/(\gamma_s+2) \tag{7.36}$$

A fair value of $\mathcal{R}LK \approx 100$ and higher is considered sufficient for the above asymptote to take place. As can be seen from the last two expressions that the threshold value is largely administered by the SNR alternations and hence an adequate adaptation policy needs to be further explored. The above derivations are given in virtue of one threshold setting only. This is commonly known in the literature as a one-stage adaptive threshold. Other studies also advocate more than one-stage threshold schemes [93].

7.6.3 Optimum Frame Duration

As mentioned earlier in the context, it is also a prime objective to optimise the minimum sensing duration to facilitate maximum achievable throughputs of CRNs without blocking PUs. This is a crucial task to allow SUs to smoothly vacate their subbands upon PUs detection in overlay CRNs. Subsequent to the derivations approach of collision duration ratio proposed in [254, 255], this task can be computed for any priori given pair of (P_{md}, P_{fa}) and complex signaling in the perspective of MTSE-MIMO. After cancelling the threshold parameter from (7.26) and solving for the minimum number of sequence points needed for the detection frame, this obtains

$$K_{min} = \frac{2}{RL{\gamma_s}^2}\left[\mathrm{Q}^{-1}(P_{fa}) - (\gamma_s + 1)\mathrm{Q}^{-1}(P_d)\right]^2 \tag{7.37}$$

It is therefore apparent that the minimum sensing interval, or OFD, is inversely related to the parameters of MTSE-MIMO (STBC) structure. That means increasing the number of antennas, tapers and second exponent of the SNR value will reduce the number of sequence samples required to achieve successful detection and higher throughputs. The inverse proportionality with respect to SNR exponent is of great impact and would be highly desirable if it can be directly done, but due to various constraints this option is not always feasible in reality.

Further to the above pioneering analysis, the IMM-KF is suggested for the first time here to estimate the SNR value and achieve the AT and OFD tasks, which is attended by using the MTSE for SS in CR applications. This constitutes yet another important contribution that has not been addressed in other studies before, and can be enlisted with other contributions along this way. The FET approach also represents a natural choice for the IMM-KF algorithm in MIMO and STBC settings.

7.7 Complexity Analysis

Few studies on the computational analysis of spectrum estimation algorithms can be recognised in the literature, most of which attempted the arrangement of single antennas. For instance, the FFT as core engine for MTSE was appraised in [257] and [258], with and without involving Slepian tapers, respectively.

Without loss of generality, applying the MTSE computational analysis could be a straightforward achievable task. However, bringing it forward serves the purpose of presentation clarity and completeness of this work. The computational complexity of the proposed MTSE-MIMO algorithm is hence assessed with respect to the conventional PSE as a baseline. The same is also exactly applicable to the MTSE-STBC case without any discrepancies. Such computation typically involves counting the numbers of all mathematical operations $\mathcal{O}(.)$ such as addition, subtraction, multiplication and division.

The expressions given in (7.20) and (7.22) are benchmark for the computational comparison between the MTSE and PSE. The main difference between the MTSE and

PSE algorithms lies in the speciality of spectrum windowing. As a base for fair comparison, various windowing techniques that lend themselves easily for the PSE can be suggested [259]. Due to its moderate intricate and performance features, the Hamming window incorporated with Welch algorithm is merited here accordingly. The frequency resolution indicated in [259] can be normalised and equated to $1/K$ for simplicity. Recall that the heavy computational demand is originated from the FFT process, which has the figure of $Klog_2K$ total combined operations. Generalizing the STBC over all operations for both the MTSE and PSE can infer the approximated demand loads shown in Table 7.1 below.

Table 7.1 Computing loads for MTSE & PSE in MIMO.

Algorithm	Total $\mathcal{O}(.)$
MTSE	$RLKlog_2K$
PSE	$RLKlog_2(5.12K)$

It is apparent from the above table that the MTSE algorithm has a computational overhead less than its PSE counterpart under the same MIMO settings. It also has lesser computations compared to other spectrum estimation methods belonging to the same category such as the cyclostationary [258]. Worth to mention is that the Slepian tapers in MTSE and arbitrary windows in PSE remain fixed once determined and their inclusion in the recursive computations is unnecessary. Also, the output of $|.|$ is a scalar and does not raise the replication of the FFT process. The outer multiplications of this scalar largely relate to the number of eigen channels and tapers, whereas the additions of which are also to be repeated in the same manner. However, the additions are of minor impact on the mass operations and hence can be fairly ignored.

7.8 Simulation Results and Discussion

In this section, numerical simulations are presented to corroborate the viability of the above theoretical analysis. The performance of the proposed MTSE using different signalling techniques, such as MIMO, STBC, and OFDM is assessed. Also the AT and OFD behaviour is evaluated with the aid of IMM-KF to provide the necessary channel

and SNR estimates. The assessment exercises are divided into two parts; theoretical to examine the mathematical equations and practical by simulating signals virtually equivalent to real scenarios. In either case, the generated results are not necessarily equivalent under each scenario for different signalling schemes since their metrics differ considerably.

1) KLT Algorithm for MTSE and PSE

In this first exercise, the assumptions are; two bandwidths of 5 MHz and 10 MHz are considered at particular centre frequencies and a sampling rate of 100 MHz. Hence, the normalised bandwidths B will be 0.05 and 0.1, respectively. The MTSE constitutes an FFT of order 128 and consequently the rough number of effective tapers will be 6 for the first channel and 12 for the second, however, selecting a few tapers between 3-to-6 is also valid since not all of them are dominant as mentioned above in the context. Setting the above parameters is essential in the design of any MTSE employing Slepian tapers.

The spectrum activity of the above two bandwidths is interrogated in the time and frequency domains. Such behaviour is highly governed by the sinc function given in (7.19), which constitutes an essential part of the KLT algorithm in case of the MTSE scheme. The Toeplitz autocorrelation matrix of 128×128 size is generated first and then the KLT applied to find the eigenvalues of this matrix. Only three dominant eigenvalues, or equivalently tapers as mentioned earlier, are selected for display. The results of both the autocorrelation matrix and tapers with respect to the normalised sequence, i.e., divide by 128 samples number, are depicted in Fig 12. Parts (a & c) of this figure show typical performance of a sinc function with the main lobe concentrated in the centre and multiple zero crossings elsewhere; Parts (b & d) of the same figure exhibit the tapers that correspond to their respective autocorrelations stated above.

Fewer number of zero crossings, and hence amplitude cyclic repetitions and decaying rate can be observed in case of 5 MHz bandwidth as shown in Fig. 7.12.a & b while the opposite occurs for 10 MHz bandwidth as given in Fig. 7.12.c & d. As expected from part Fig. 7.12.b & d of these graphs, the tapers' amplitudes exhibit a sinusoidal like trend. As the order of tapers increases, the cyclic repetition of these amplitudes also increases, while maintaining the same maximum and minimum amplitudes. An exception to that can be visualised for the first taper, which has

maximum amplitude and without a minimum one since it has one different positive cycle than the other tapers.

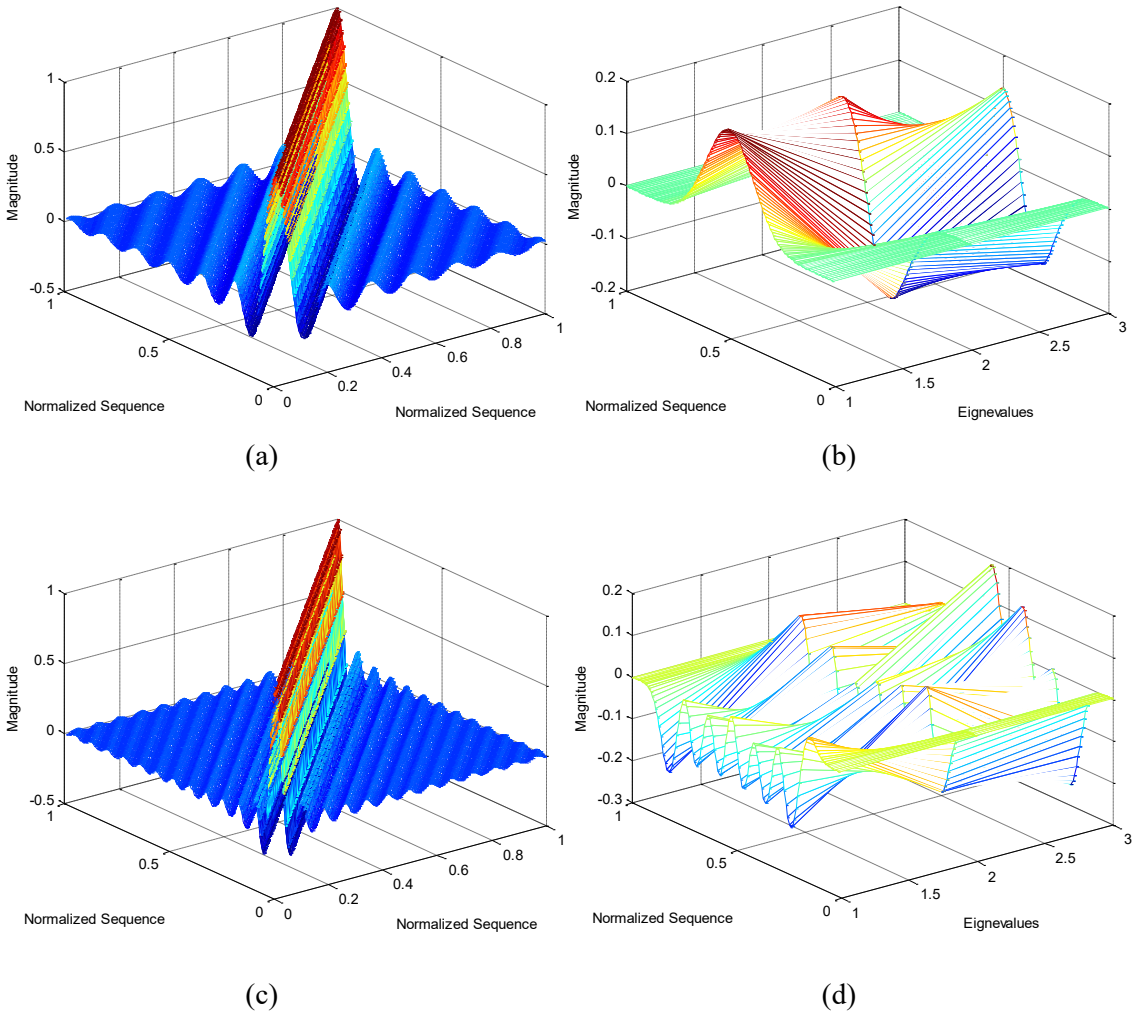


Fig. 7.12 MTSE magnitude time response (a) autocorrelation for bandwidth 5 MHz (b) first 3 Slepian tapers for bandwidth 5 MHz (c) autocorrelation for bandwidth 10 MHz (d) first 3 Slepian tapers for bandwidth 10 MHz

The focus is now directed toward the tapers behaviour in the frequency domain. Applying the FFT process on tapers will generate the graphs displayed in Fig. 7.13 below. It is very obvious that energy is densely concentrated with almost constant gain in the main lobes around particular centre frequencies, while a minor portion of the remaining OOB energy is fast fading and dispersed over side lobes. Such OOB energy experiences sharp cutoffs controlled by the design bandwidths of 5 MHz and 10 MHz. In both cases, and also contemplated for different design bandwidths, it can be

conclusively seen that the energy of the main band does not tolerate any OOB leakage penetrations. This hence would facilitate robust leakage controllability that is highly favourable in a wide range of SE applications, and attributed to the involvement of Slepian tapers.

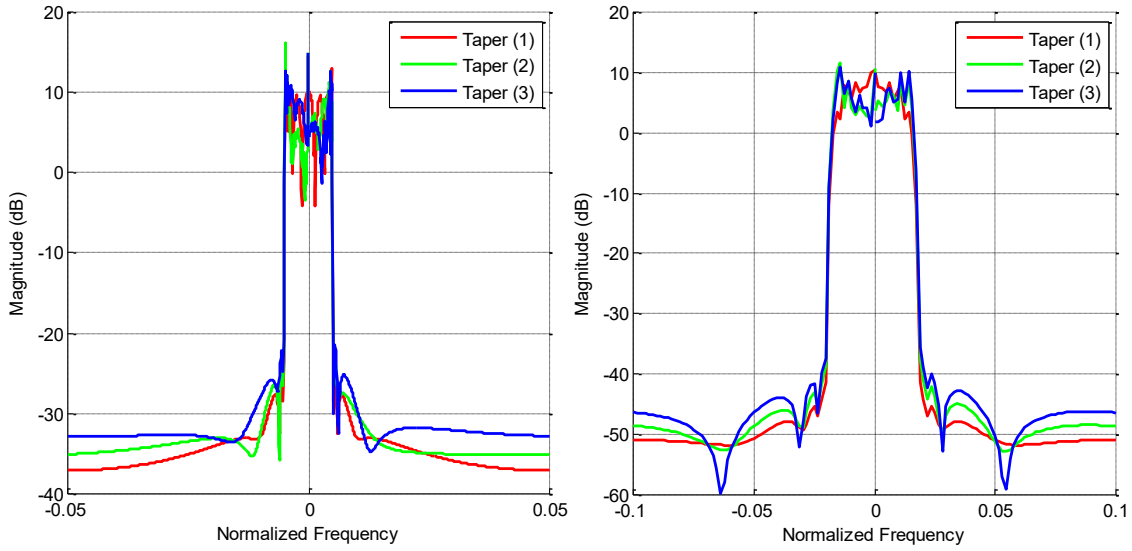
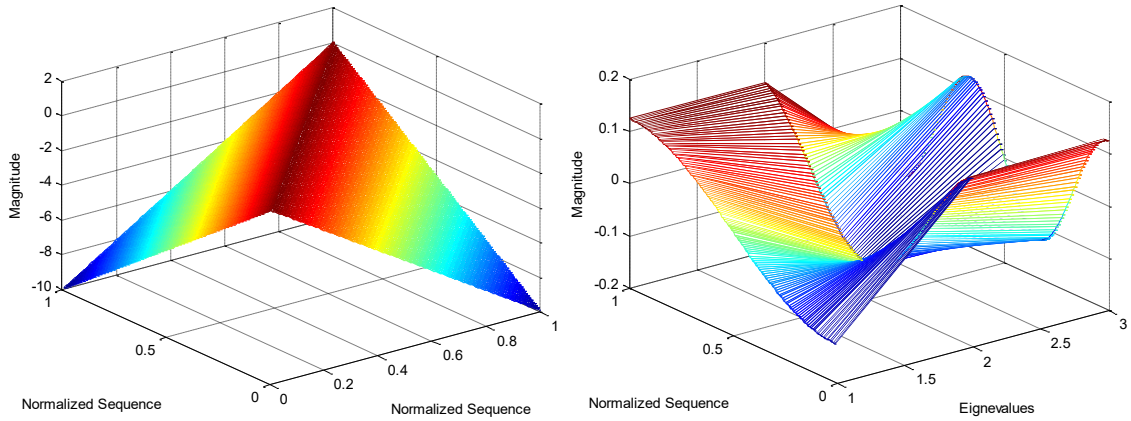


Fig. 7.13 MTSE magnitude frequency response of first 3 Slepian tapers (a) B= 5 MHz (b) B= 10 MHz.

If the attention is turned toward the PSE, a rather different version of this situation can be seen. Firstly, by examining the autocorrelation and eigenvalues, they do not change their behaviour with respect to bandwidth variations. Fig. 7.14.a represents a triangular shape autocorrelation, which is typical for rectangular windows, while Fig. 7.14.b displays their pertinent eigenvalues, and both are normalised over the time scales. The triangular autocorrelation functions are induced due to employing pure rectangular windows of low performance. No matter how design bandwidths vary, the triangular functions will not produce zero crossings and hence cannot offer further controllability over spectrum leakage. However, their eigenvalues' counterparts show no identical values and yet the sinusoidal trend can still be distinguishably recognised. Irrespective of bandwidth settings, the eigenvalues, or tapers, manifest more like sinusoidal patterns governed by the particular autocorrelation functions at hand.



(a) (b)

Fig. 7.14 PSE magnitude time response for B= 5 MHz and 10 MHz (a) autocorrelation (b) first 3 eigenvalues.

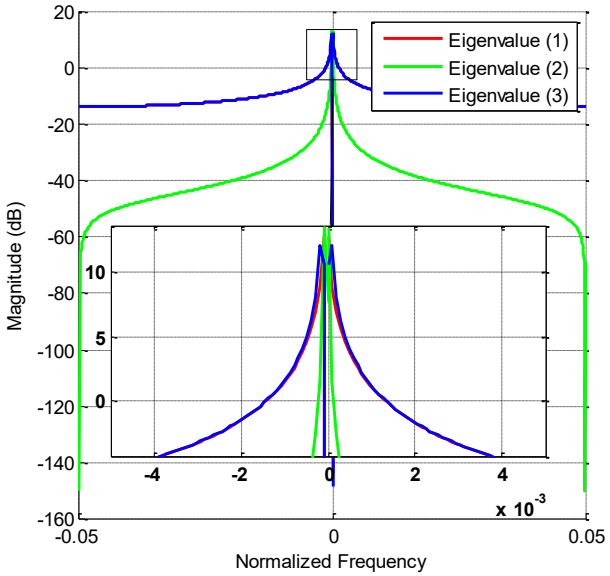


Fig. 7.15 Magnitude frequency response of the first 3 eigenvalues in PSE for both 5 MHz and 10 MHz.

More into the PSE performance, the frequency response of its underlying eigenvalues is now inspected. As revealed in Fig. 7.15, it is fairly evident that similar responses obtained regardless of the design bandwidth values. The energy in the main lobe is dispersed and does not seem to be focused around centre frequencies anymore. An overlap occurs between main and side lobes and distinctive boarders between them cannot be identified easily. More energy is spilled over a wide range of side lobes without featuring any controllability to constrain them inside the main lobes. In addition to that, the odd ordered tapers seem to follow almost exactly similar trends and this is

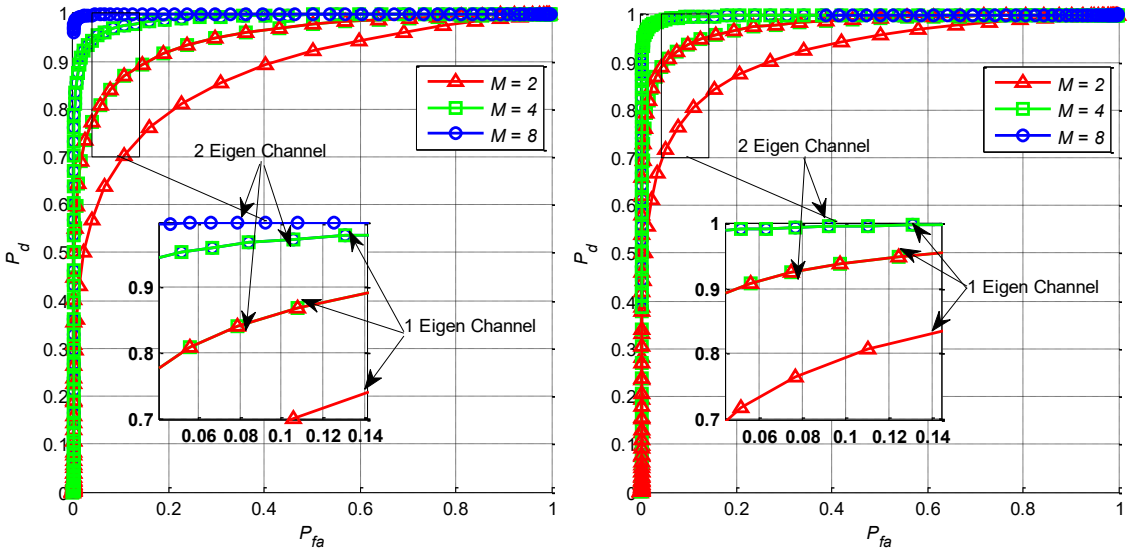
why a magnifier tool is needed to make them clearer. For example, the first taper lies behind the third taper and so forth for other tapers, as shown in Fig. 7.15. Also, evenly ordered tapers demonstrate rapidly declining trends while kicking away from centre bands. Generally, such unmanageable tapers' behaviour would render the PSE very poorly and the overall performance of which unreliable. Owing to the given analytical valuations, the viability of having the MTSE paramount to PSE can be easily corroborated in this study.

2) ROC for MTSE and PSE

Now, the multiple antennas formation effects on the proposed SE performance are to be examined. Diversity schemes with 1 and 2 antennas arrangement over Rayleigh fading channels are considered. The ROC curve represents the relationship between the two probabilities pair (7.26) for MIMO and (7.27) for STBC, respectively. The ROC has a priority preference in the assessment of the overall system performance and hence must be investigated. Two ROC numerical examples are computed as depicted below in Fig. 16.a & b, respectively. The dominant taper of values 2 and 3 differentiates between these two examples, while other settings remain common. Channels of 1 and 2 parallel eigen paths are employed. The constellation size is varied from 2, 4, to 16, while the STBC code length is simply kept to 2 for ease of computing and generating discernible results. The SNR value is set to 0 dB, while the threshold values are controllable within a feasible range. As expected, the performance of MTSE-STBC gets better boosts as the number of eigen channels increases. Further improvement can still be observed when the M-ary constellation size increases.

Fig. 7.16.b clearly demonstrates that the more tapers are involved, the better the performance. While such a trend might look appealing from the outer set, however, it is partially true in real implementations. It should be emphasised that increasing the tapers' size without proper limit will unnecessarily elevate the computational burdens without extra tangible gains. It is also worth noting that the left corner on top of each figure mentioned above shows acute rushes in the performance curves, which are mostly appreciated despite not easily distinct. Such graphical behaviour is typical in situations where two or more performance trends are combined with high pitch differences on the same metric scales. Another fact related to these graphs is related to the range of threshold variations. In both graphs, the threshold values have been

changed in the same range and this is a correct approach in order to assess the sensitivity against the changes in other parameters on the detection resolution. The detection resolution is improved by increasing the number of the parallel eigen channels, M-ary and taper orders as anticipated.



(a) (b)

Fig. 7.16 ROC of MTSE-STBC against M-ary variations (a) 2 tapers (b) 3 tapers.

Another useful assessment tool is to plot the ROC with respect to variations in the SNR value. Such a chart is depicted below in Fig. 7.17 for different SNR values (-1 dB, 0 dB and 1 dB), while using a few numbers of received data samples, 2 eigen channels and 2-ary signalling. The number of tapers is 2 in Fig. 7.17.a and 3 in Fig. 7.17.b. The given trends confirm a better performance gained with respect to an increase in the SNR and taper orders, which is naturally expected. It is important to realise the fact that the dynamic range of SNR variations could be very large in real situations. Also recall that the overall MTSE performance considerably pitches higher by employing the STBC and MIMO techniques, which is the core intent of the analysis of this work. This is why such a small range of SNR variations is selected, as given above in order to maintain a clearer vision of the ROC trends. This also justifies the tendency of the ROC trends to shift toward the top left corner of the ROC figures as the SNR increases dramatically, or even by increasing the number of tapers, constellation and spatiotemporal coding. The

same ROC arguments apply equally to the MIMO situation in the absence of the direct coding improvement contribution.

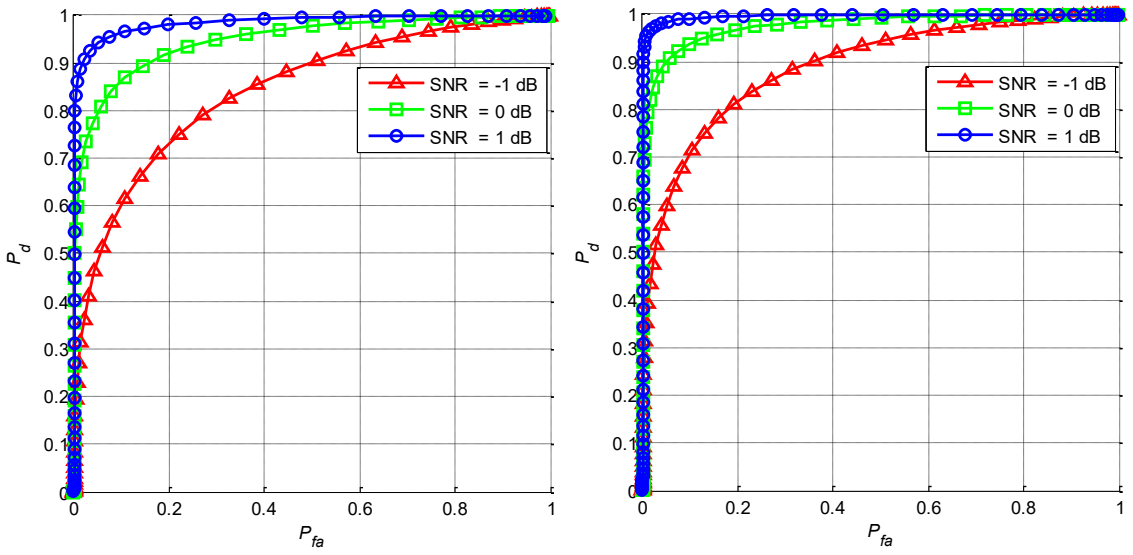


Fig. 7.17 ROC of MTSE-STBC against SNR variations (a) 2 tapers (b) 3 tapers.

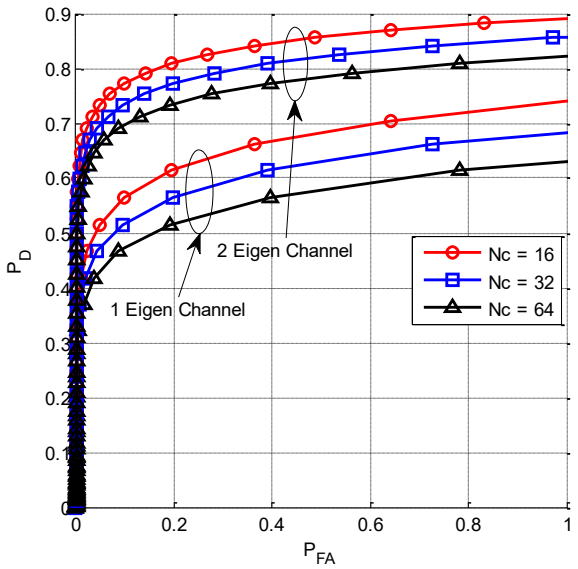


Fig. 7.18 ROC of MTSE with 2 tapers and MIMO-OFDM versus FFT and channel orders.

The final exercise in this group is about applying the MTSE using MIMO-OFDM signalling. Aiming at computational ease and having clear graphics as far as possible, two dominant tapers are used with 1 and 2 parallel paths channel, respectively. The M-

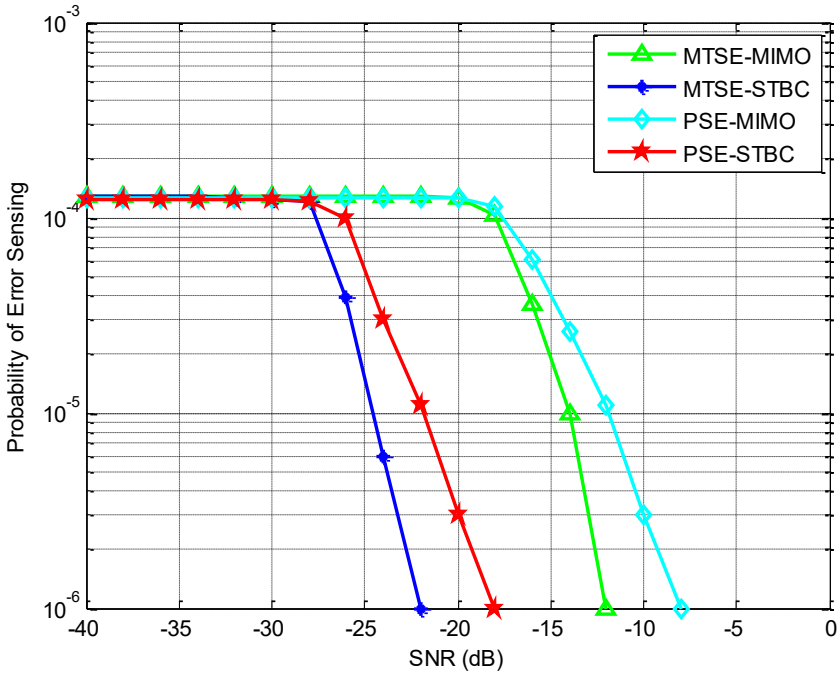
ary size has been intentionally fixed to $M = 2$ to manifest the case of BPSK. The number of subcarriers varied from 16 to 64 and for several data symbols and the range of subcarriers equals the FFT order in this case. Fig. 7.18 above illustrates the first ROC exercise employing 2 tapers. As expected, despite the ROC experiencing a better trend when the number of dominant channels is increased, however, an adversary performance is obvious with respect to increasing the FFT order. The larger the FFT order, the worse performance is obtained and this confirms that the OFDM has more OOB spectrum spillage and hence does perform so well for SS applications in CRSs. This also agrees with the findings of other earlier studies in this regard.

3) Simulated Performance of MTSE and PSE

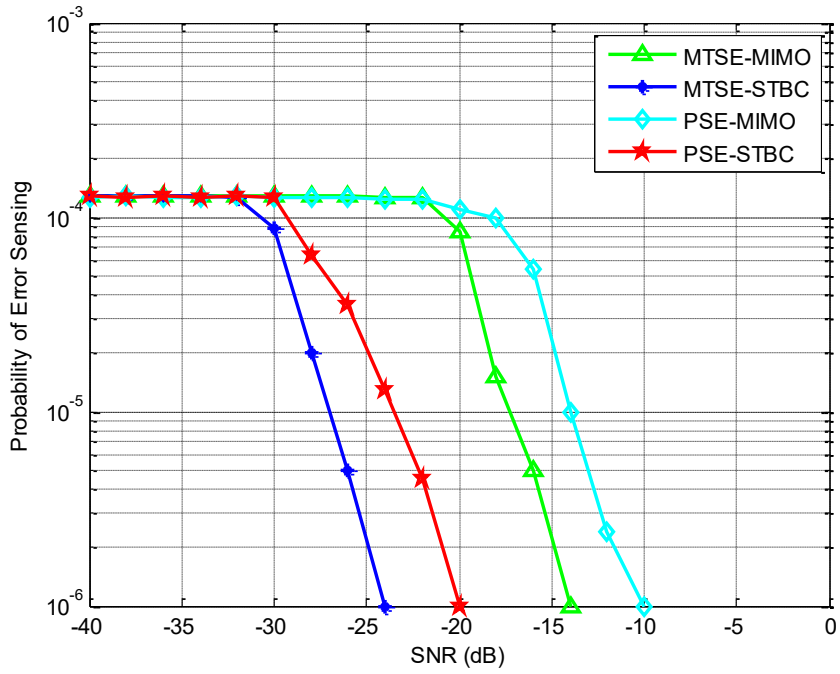
This exercise is concerned with the simulation of a BPSK random sequence transmitted over a Rayleigh channel contaminated with AWGN. The number of samples chosen here coincides with the moderate range between 50 to 200 samples, and sometimes as small as 15 or 20, to satisfy the SE functional requirements [90, 237, 244]. Regarding the number of samples, more than 100 was shown to be enough to tackle the faded signals and spread spectrum issues and adequately attain fast computations [237]. Two configurations of 2×2 MIMO and Alamouti-STBC are utilised. The data sequence is interrogated using PSE and MTSE schemes with a prefixed threshold value tantamount to CFAR of 10^{-3}. The MRC is applied on the MIMO receiver side to maximise the SNR.

A Monte Carlo of 10^5 repetitions is conducted and the resulting probability of error sensing against SNR is depicted in Fig. 7.19.a below. While the multi-antenna settings are generally well cherished for their positive impact, however, the first conspicuous glimpse on this figure reveals the competent performance of STBC compared to MIMO. The chart turning point is around -15 dB for MIMO which is nearly typical in other ample studies, while it is more advanced at about -25 dB for STBC.

The same example is repeated for 4×4 multi antennas while maintaining other design settings. The signal is altered to be of 4-ary PSK or QAM and the square generator matrix of Alamouti code is expanded to have 4×4 size and maintain the coding rate of one. The MTSE and PSE results of such a scenario are depicted in Fig. 7.19.b below.



(a)



(b)

Fig. 7.19 Probability of error sensing against SNR (a) 2×2 antennas (b) 4×4 antennas.

As can be seen from this figure, the turning points are displaced by a certain amount for both the STBC and MIMO probability of error detection trends. It is now roughly near -17 dB for the MIMO, while it is around -30 dB for the STBC. The major improvement in the STBC performance is attributed to the increase in the coding gain,

which is about 3 dB more than the case of 2×2 equally applicable to both the PSE and MTSE. The constructive contribution of space diversity is also adding up to the overall performance gains displayed in this figure. But still the shear difference between the STBC and MIMO situation remains very distinct and in favour of the former. Increasing the space diversity has a minor impact on the MIMO responses, while it is of preferable influence on those belonging to the STBC. Putting this in an appropriate perspective, this is mainly due to more dominant eigenvalues contributing to such major performance boost in the case of STBC compared to MIMO.

Another main advantage linked to the STBC can be distinguished by closely examining the detection thresholds given in (7.35) and the related explanation above. Since the threshold value has been set to accommodate the STBC, the MIMO thereby roughly needs a minimum of M times the power of STBC to climb the hill beyond such a threshold. Therefore, the overall performance gap between the STBC and MIMO systems is in general gained by the compound contribution of both advantageous factors stated above, which is about $20log_{10}M$ warranting the STBC system. Such behaviour is in total agreement with the analysis and descriptions highlighted in the above context.

As stated above, the STBC enjoys additional eigenvalue gains and robust spectrum leakage management. Thereby and under both transmission scenarios, the MTSE leaves no doubt about its performance superiority compared to PSE. Such outstanding MTSE performance is mainly attributed to the excellent spectrum control policy achieved by Slepian tapers. Appendix (B) illustrates the pseudocode for the MTSE and PSE algorithms.

4) AT and OFD

Assume a BPSK sequence of K=1024 samples applied to MTSE with L=3 Slepian tapers and equipped with 2×2 MIMO configuration and through flat-fading Rayleigh channel. The total number of effective points within the processing frame is equal to 6144 in this case, which surely rises as the complexity of MTSE-MIMO increases while maintaining the signal sequence intact. Therefore, the asymptotic approximation (7.36) is evidently met.

The performance of detection threshold is firstly assessed with reference to SNR variations from -25 dB to 0 dB, and arbitrary utilisation factors α = 0.1, 0.5 and 0.9.

The value α = 0.5 symbolises the separation line unravelling two preferences maintaining; better spectrum reuse (larger α) or additional PU protection (smaller α). Three threshold trends are thus generated as per expression (7.35), while the fourth trend represents the asymptotic threshold behaviour described in (7.36) and all are plotted in Fig. 7.20.

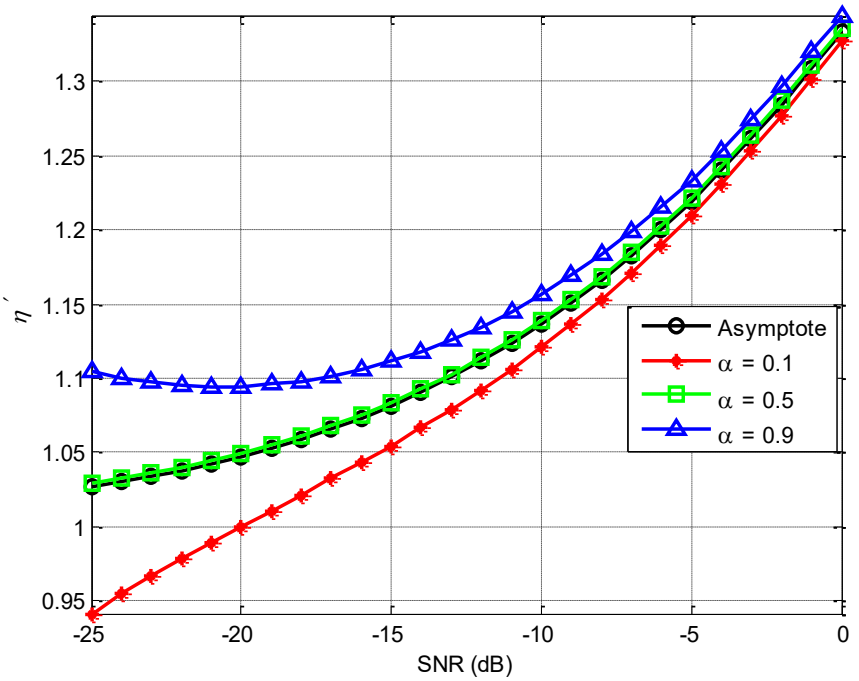


Fig. 7.20 Detection threshold against variations in SNR γ: asymptote and selected values of utilisation factor α.

Fig. 7.20 infers several observations that are worth looking at. 1) The asymptotic threshold trend is in exact alignment with the threshold trend of α= 0.5. 2) All threshold trends are concentrated and sharply climbing with respect to the incremental SNR beyond 0 dB and irrespective of α. This means the received SNR value is relatively large enough to readily bring the requirement of appropriate spectrum reuse and PUs protection correspondingly applicable. 3) The threshold trends for $\alpha < 0.5$ or $\alpha > 0.5$ exhibit a symmetrical distribution cantered at the asymptotic and α = 0.5 threshold trends. 4) The threshold trends for $\alpha >$0.5 are also mounting for SNR less than 0 dB. This seems logical to set for larger threshold values in order to secure better PUs' protection designated by larger values of α. Keeping in mind that the utilisation factor of (7.32) and (7.35) in this work is allocated different to what was given in (Eqs. (6) and

(11) of [249]). However, the compilation of (7.32) and (7.35) is consistent with that given in (Eqs. (9) and (12) of [250]). Therefore, the outcomes illustrated in this work are intuitively assumed to be more persuasive compared to other findings given elsewhere. Whatever the case is, an AT policy consequently needs to be appropriately developed to address SNR variations. Such a task can be administered by having the results of (7.35) substituted in (7.33) for predefined probability metrics and the active spectrum reuse can be attained as in (7.32) accordingly.

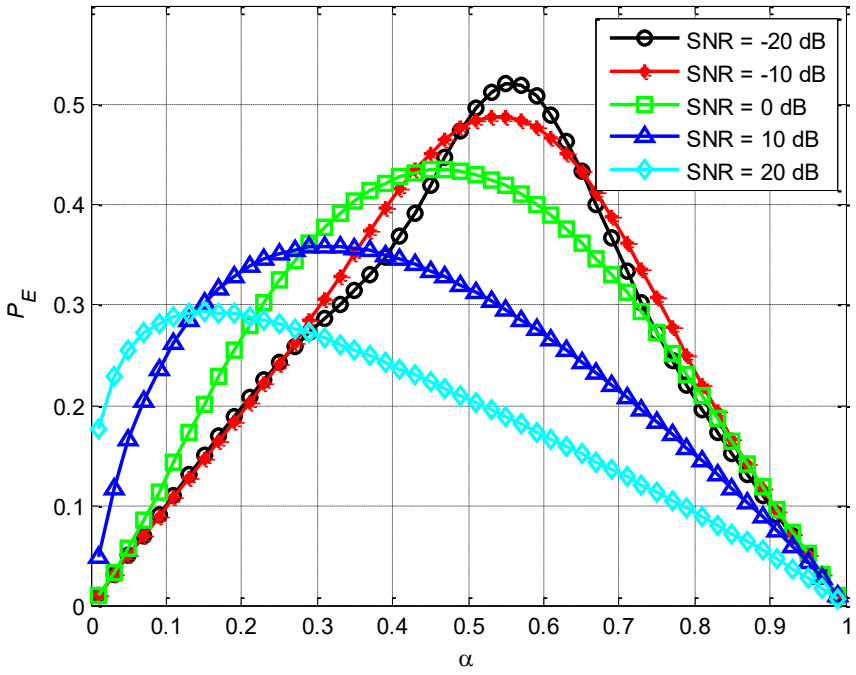


Fig. 7.21 P_E against variations in utilisation factor α.

The next simulation exercise is on examining the total error decision probability P_E against variations of α values in the range 0 to 1, and for wider range of arbitrary SNR values -20 dB, -10 dB, 0 dB, 10 dB and 20 dB, as depicted in Fig. 7.21. The noise variance is σ_n^2 = 1 and the prefixed threshold value fixed to have CFAR or P_{fa} of 0.01. This value is commonly used in line with the maximum acceptable figure of 0.1 for both the P_{fa} and P_{md} defined in the IEEE 802.22 regulation [90]. A higher P_E means lower spectrum efficiency for both the PUs and SUs and vice versa is true. Therefore, P_E needs to be minimised to the best extent as possible to have enhanced spectrum performance. Fig. 7.21 shows that the worst case scenario for P_E occurs around α=0.5, especially for lower SNR values. That means there is no preference of distinguishing

between SUs or PUs while accessing and using particular bands. On the other hand, the best P_E performance, and thus the spectrum utilisation, can be achieved on either side of the depicted convex curves. That is lower α values can be assigned for either function of having better SUs' spectrum utilisation or higher α values for PUs to enjoy better protection against opportunistic bands access. The shifting of such convex shapes towards the left side of Fig. 7.21 for higher SNR is also obvious. This is attributed to the enhanced signal and channel condition that allows for better PUs detection under almost a very wide range of α values much below 0.5. Therefore, under such higher SNR range above 0 dB, a better spectrum utilisation and PUs protection is almost always achievable.

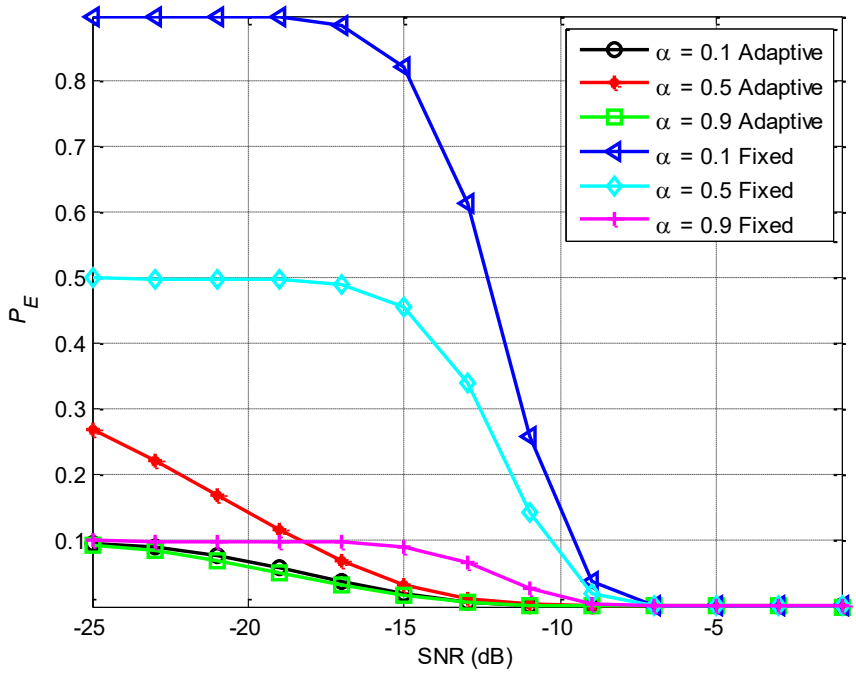


Fig. 7.22 P_E against variations in in SNR γ.

The next exercise is devoted to examining the proposed adaptive and the classical fixed P_E performances against SNR variations and with respect to arbitrary utilisation factors, as shown in Fig.7.22. It shows that worst case performance, or higher P_E, occurs at α= 0.1 or far below 0.5 and particularly at smaller SNR ranges. At such ranges, the only way to improve the P_E performance is by making the utilisation factor much higher than 0.5, such as α= 0.9 or above. The P_E for α= 0.9 shows a much

improved performance as shown in the same figure. The first scenario means no solid protection for PUs, while the second indicates exactly the opposite. As for α= 0.5, the P_E is obviously exhibiting poor performance and the trend of which just lies in the centre between other trends for α= 0.1 and 0.9.

As that is said on the fixed threshold, the AT on the other hand enhances the overall P_E performance to significant levels. This is true in as much as α takes on values either smaller than or larger than 0.5. Again, the P_E performance for α= 0.5 is modest; however, it is still much better than the fixed threshold situation, as shown in Fig.7.22. A precaution needs always be considered whether to make further PUs protection or allow SUs to flexibly access and use vacant spectrum without paying full attention to PUs' activities in the same or adjacent bands. This can only be decided by allocating a proper utilisation factor α, either much greater or less than 0.5 for any of the above situations, respectively.

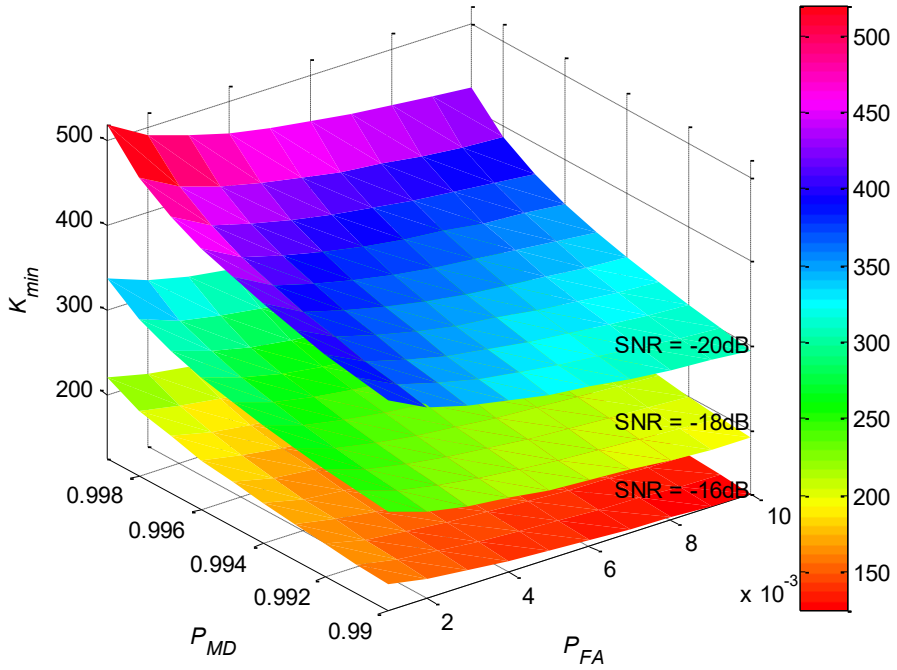


Fig. 7.23 Minimum sensing duration against variations in (P_{md}, P_{fa}) and selected SNR γ values.

The last exercise in this group is around demonstrating the minimum sequence samples required for SS to meet the permissible optimum frame duration. Fig. 7.23

shows that the minimum samples condition constitutes nonlinear hyperplanes of two arguments, namely the pair (P_{md}, P_{fa}). Multiple non-overlapping hyperplanes are also obvious for different SNR values. Three arbitrary SNR values of -20 dB, -18 dB and -16 dB, typical in SS environment are assumed. The IEEE 802.22 mandated ranges for CFAR and CDR values are conformed in this example. As predictably exhibited in Fig. 7.23, the hyperplanes foci are greatly reduced relevant to the increase in SNR values and hence less sample points can realise the minimum of optimum detection frames. On the other hand, larger sample extents are entailed for utterly deteriorated SNR conditions. The minimum sample points are displayed on the K_{min} axis of Fig. 7.23, which signify the peaks of hyperplanes computed for the targeted lesser values of (P_{md}, P_{fa}) pair.

As for the design pair (P_{md}, P_{fa}), smaller setting values entail fairly larger number of sample points to accommodate efficient SS application, especially under inferior SNR circumstances. Recall from the experimental settings of L=3 Slepian tapers and equipped with 2×2 MIMO stated earlier in this section, it becomes well discerned that the value of K=1024 can easily achieve any targeted pair (P_{md}, P_{fa}) shown in Fig. 7.23. Hence, it is always advisable to engage an acceptable range of sample points that can attain the minimum OFD demand without further complexity overheads.

5) AT and OFD Using IMM-KF

This is the last assessment exercise, which is concerned about applying the IMM-KF algorithm to adapt the AT and OFD under realistic signal simulations. The same simulation settings used to generate the results in Fig. 5.11 and Fig. 5.12 are repeated here. The IMM-KF in this case is employed to estimate the SNR and insert its value in (7.35) and (7.37) to determine the AT and OFD, respectively. The number of tapers and utilisation factor are set to 3 and 0.5, respectively. The latter means PUs and SUs are equally evaluated. Fig. 7.24 below shows the general AT and OFD trends, as normalisation is applied to the first. As expected, the AT level increases as the SNR rise, while the opposite occurs for the OFD and hence lesser number of samples are required for detection. This is mainly attributed to the healthy estimate of SNR attained by the IMM-KF and employed to reconfigure the AT and OFD accordingly. As Fig. 7.24 below shows the instantaneous trends, averaging of the same parameters could be advantageously recommended for use. The number of samples for detection can also be

inversely reduced with respect to Slepian tapers and the MTSE can thus be tasked for better throughput.

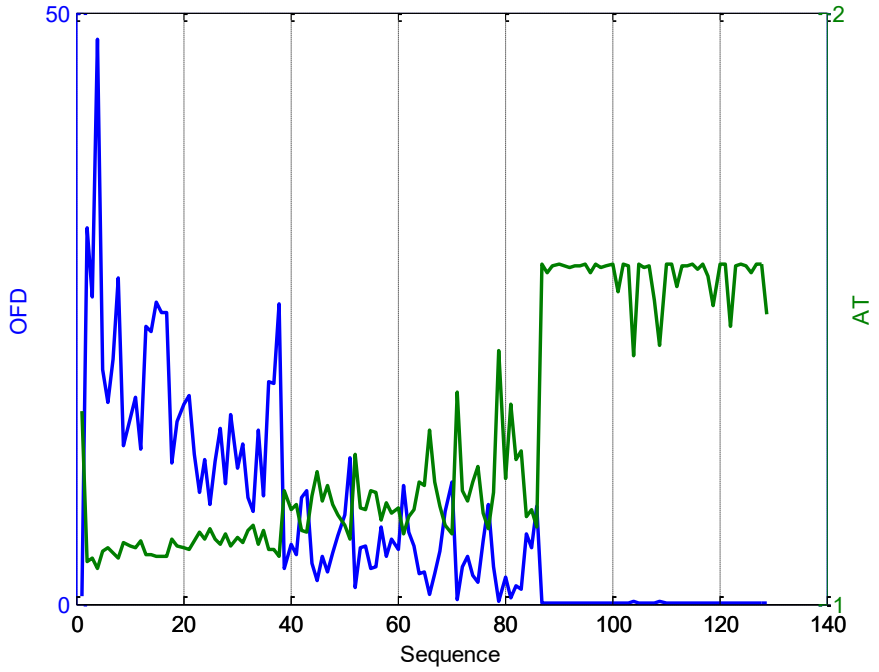


Fig. 7.24 Estimated AT and OFD.

7.9 Summary

This chapter has discussed the SS application, which is a pinnacle among others in the premise of CR systems and CRNs. Most strategies related to the efficient spectrum reuse, access and management heavily rely on resilient and reliable SS methodologies. As such, the MTSE has been presented to seek excellence in satisfying such rigorous objectives. The non-parametric SE using the MTSE and the PSE schemes have been developed. The blind approach has been adopted in the analysis and design of these two schemes in as much as the unknown parameters are not given. However, the non-blind approach can also be implemented in the same exact manner, but in this case some sort of CSI estimations like the MLE, or any other estimator, must be employed. The known or estimated CSI can help in the formation of NP decision policy to process the PSD as a sufficient test statistic. Upon the Chi-square test of PSD statistic exceed certain

threshold values infer on the PUs' presence, otherwise, PUs are not available and their particular bands can be used by SUs freely.

The two performance metrics; P_{fa} based on CFAR and P_d on CDR, have also been formulated for the MTSE under certain signaling arrangements. These two metrics usually rely on predefined CFAR or CDR values and based on which the MTSE performance for SS can be evaluated. Each metric is used to quantify a particular threshold value that the SS apparatus needs to comply with eventually. Focusing on the CFAR offers more protection for PUs, while the opposite is true for the CDR and hence an appropriate compromise needs to be worked out in light of the given operational conditions.

Based on the above, some sort of utilisation factor has been embedded in the construction of the overall P_E formula to balance the weights assigned for both P_{fa} and P_d. The values of such utilisation range from 0 to 1, where lower values provide SUs with better spectrum opportunities while higher values give PUs better protection. Based on that, an arbitrary threshold adjustment is to be wisely devised to cope with the changing situations. An adaptation policy has thus been elaborated to produce dynamic AT and OFD as per the requirements drafted in the IEEE 802.22 standard. The IMM-KF was featured as an efficient CSI estimation structure to reliably feed with certain CSI estimated parameters entailed to perform the AT and OFD adjustment properly.

Simulation methods have been conducted and revealed the supremacy of the adopted MTSE for SS under a wide range of operational conditions. The results have also shown the SS performance enhancement is conceivable in the cases of MIMO and STBC schemes implemented. The latter has better performance than the first. While, on the other hand, the OFDM has shown inferior performance and hence not so welcomed in SS applications. Moreover, the performance of the proposed MTSE has been proved to be much better compared to the PSE technique under all conditions. Furthermore, the IMM-KF, once again, proved itself as a very reliable CSI estimation tool and can be flexibly used to adapt the various SS parameters in order to achieve better spectrum reuse, admission and exit as per the prevailing policies. Therefore, this chapter has been well perceived in delivering several new paradigms that were not fully addressed elsewhere.

Chapter 8

CONCLUSIONS AND FUTURE WORK

"A journey of a thousand miles begins with a single step".

By the Chinese Philosopher Laozi

8.1 Concluding Remarks

The journey of this study has diversified the wireless communication techniques that constitute essential foundations for the CR systems. The prevailing wireless techniques adopted in well recognised standards such as IEEE 802.11, IEEE 802.15, IEEE 802.16, IEEE 802.20, and IEEE 802.22 are considered as corner stones in developing and building CR systems. They are based on the envisioned SDR and are well suited to work under various environments and working conditions. As such could be interpreted as friendly environment and hence hand shaking can be implemented between transmitter and receiver, otherwise communication patterns recognition need to be considered under hostile conditions. Above, the call for sustainable spectrum access and reuse can be smartly attended by the CR systems. Along this way, a CR platform calls upon a variety of smart gears to be able to cope with different work environments. The tools presented in the forgoing chapters are quite essential for CR systems to help combat severely fluctuating and fading channels in the presence of PUs and SUs aiming to occupy licensed or unlicensed spectrum holes at any expense. CR systems enjoy multiple features and hence are anticipated to be a hot research area for decades to come due to their presence in vast applications. Emergency, search and rescue, entertainment, multimedia, directions and people finding are just a few among many interesting services that can be enumerated about CR capabilities. Updating services using online methods is also anticipated feasible by simply sending appropriate request to the spectrum controller in the same manner as the Internet does. The drive momentum for CR systems is to explore and develop a wide range of human oriented applications is predicted with high potential. The current trends of wireless

communication systems and behind them the CR methodologies are entering new paths and promise for a bright future. Great efforts are underway to shape up mankind's future towards directions of more intelligence and be smarter. CR systems are well versed into such paths and conform better to numerous needs. It can be loudly voiced, and without any hesitation, that CR systems will open an immense horizon to new applications that were never imagined feasible before, thanks to the fusion between the rapidly evolved nanotechnologies that has paved the road for well integrated software and hardware platforms. This study is believed to be of a significant value in shedding suitable light on the CR development opportunities. Also it contributes to pave optional routes of ample potentials to achieve higher performance levels projected for future smart wireless communication systems.

The following are summary remarks inferred from each chapter of this work.

Chapter Two: A review of the SDR and CR concepts and the technologies along their most important milestones' evolution has been presented. The most important standards involving the main CR functions, in their current or developing schemes, and their domains of speciality have been described. The state-of-the-art studies focusing on the various aspects of CR systems and networks with their potential applications have been brought forward. Overviews of the main three functions that constitute the backbone of any CR platform, namely the ACM, AMC and SS have been thoroughly exemplified. Some of the multifaceted operational aspects that embrace one or all of the aforementioned prime CR functions under diverse transmission settings have well been described as well. A reader of this chapter is expected to gain a meaningful idea on the roots, meanings, development efforts that are necessary to advance into further explorations in the future.

Chapter Three: The tools for prevailing digital signalling and coding techniques that are essential to fulfil the CR transmission requirements over typical wireless communication channels have been provided. For example, baseband M-ary digital signal (PSK, QAM) and system modelling, multi antennas and carrier, spatial and time coding (MIMO, STBC, OFDM), in addition to the gains achieved by applying the various diversity and combining techniques have been well demonstrated. The performance metrics of all of the aforesaid signalling and transmission settings have been analysed and then evaluated by using simulation methods.

Adding to that, Chapter Three has also tackled the most appropriate modelling schemes for Rayleigh fading channels, namely; the Gilbert-Elliot, Jake and Clark, and the famous FSMC. These all have been developed using appropriate AR and fading states' orders to reflect on the desired dynamical range of studied variations. Probably the most important contribution above all in this chapter is to borrow a new concept commonly used in the econometric studies and projected on to the FSMC model. This innovative paradigm is based on the information-theoretic approach using Tauchen model, which has been tailored for the FSMC modelling of Rayleigh fading channels for the first time in this work. Simulation results corroborated the improved suitability of the new paradigm of Tauchen in modelling Rayleigh fading channels based on the FSMC approach.

Chapter Four: The elemental foundations of estimation theory and then the classical CSE techniques, such as; LS and MMSE, have been investigated. The CSI estimation is necessary in various blind applications when there is no advance knowledge provided on the CSI or there is no cooperation between transmitters and receivers. The state-space estimation technique exploiting the reputed KF has been elaborated. To enhance the robustness of KF and make it more adaptive to a channel's variations, the IMM algorithm has been augmented to produce the IMM-KF scheme. As a novel contribution, the FET approach, which depends on the sum of effective eigenvalues of the disintegrated channels by using SVD, has been proposed. By applying the FET on the classical and the IMM-KF merger, the later has shown promising results in estimating and tracking the CSI under Rayleigh block or flat fading channels. Therefore, such desirable IMM-KF behaviour has inspired its further embracement as a consistent driver for reliable CSI estimation in the remaining chapters.

Chapter Five: The ACM, as a prime function among the other two in CR systems, has been illustrated. The need of having an appropriate feedback channel between the transmitter and the received under friendly or cooperative work conditions has been thoroughly elaborated. This is required in order for the transceiver system to achieve the best spectrum utilisation under various varying work conditions. And here the IMM-KF suggests itself as a powerful candidate to stratify such requirements and has been embedded in the ACM structure accordingly. Such a novel paradigm has been contributed for the first time in this work. The discrete VRVP and the WF concepts as

common ACM techniques have been analysed. The channel coding techniques chiefly represented by the BICM and TCM have also been examined. Simulation results have corroborated the given analysis, and that the BICM is well suited than the TCM for fading channels and the IMM-KF is a very good candidate for link adaptation and achievable spectral efficiency, which is worth further explorations.

Chapter Six: The AMC as a second pillar in the CR systems has been analysed. Unlike coherent detection of friendly systems, there is no harmony between transmitters and receivers under a hostile environment. Therefore, the blind identification of M-ary PSK and QAM alphabets in the context of likelihood functions has been explored accordingly. The pressuring demand for the JCEDD orchestration has been clearly pronounced to help achieve better AMC in blindness. In addition to the proposed IMM-KF algorithm, the MLE has been tasked to fulfil such requirement. The EM interlaced loops, among other approaches, have constituted an essential part in the realisation of the JCEDD design, which is based on IMM-KF amalgamation. The complexities of both algorithms have been worked out and simulation results have confirmed the suitability of IMM-KF to achieve the optimum MLE performance with less computational overheads.

Chapter Seven: This chapter has been devoted to study the most important functional objective of CR systems, which is expressively stressed to be the SS alongside the aforesaid critical ones. The classical FBk, periodogram and MTM for spectrum estimation have been analysed. The proposed MTSE has been furtherly elaborated due to its robust performance since the energy of which is concentrated in the centre of the bands of interest with least OOB leakage. That is mainly attributed to its employment of very powerful spectrum truncation tools called Slepian tapers. In the premise of Bayesian decision statistics, the NP has been featured to maintain the decision making on the PSD as an indication on the presence or absence of PUs. Eventually, SUs can efficiently detect empty spectrum slices and opportunistically use without generating any harm to PUs. Also, SUs should vacate these slices upon the request of PUs and within the times specified in the related standards such as the IEEE 802.22.

The adaptive thresholding and optimum frame duration, namely; AT and OFD, which are essential to accomplish the best SS strategies, have also been studied in

Chapter Seven. That is in the sense that the probabilities of correct detection and false alarm, P_d and P_{fa}, respectively, are heavily reliant on the prefixed metrics of targeted SS performance. Such metrics can be commonly determined by either using the CFAR or the CDR approach. The first prepares for better SUs spectrum access, while the second pushes for more PUs protection. Therefore, a tradeoff by using an arbitrary utilisation factor α is facilitated to maintain a balance between the two compelling requirements. The IMM-KF scheme has been proposed in the design of these MTSE attributions as it has never been addressed before and hence considered as a new paradigm as well. The SNR estimate delivered by the IMM-KF has been inserted in the reckoning of the above thresholding and frame structures and analysed properly. The AT and OFD trends in the computational exercises have shown a great agreement with the results of other studies, which leaves no doubt that the IMM-KF is quite capable to achieve the intended SS tasks professionally.

Further analysis and results have also been conducted under different transmission settings, such as the MIMO, STBC and OFDM, in this chapter. In all the simulation results, the MTSE has definitely shown a superior performance compared to that of classical PSE. Moreover, and as expected, the STBC performance has shown better than MIMO, while the OFDM performed poorly. The OFDM can hence be concluded as a bad transmission medium for the SS to take place, unless other alternatives sought such as FB multicarrier systems.

8.2 Future Work

The interest of this study is motivated by setting certain goals to attend multiple development trends forecasted for the CR system and CRNs in relevance of the mankind trending in the present and the near future. The structure of future work plan will extend the trends established in this research, leading to detailed design strategies for the CR systems and CRNs and their operational characteristics under various scenarios. The following points have been recognised as key transformational potentials for future development:

Cooperative CR systems and CRNs: The design and development of the systematic trends presented in this study can be furtherly extended to explore the effects

of MU arrangement in terms of spatiotemporal attributes. PUs and SUs can be sporadic in the time and the location constraints, and their data can be either centralised or fused to gain more insight on the channels' status, and therefore worth to be examined. That is expected to enhance the spectrum sustainability and open wide doors for unexplored application areas such as emergency calls, rescue, and disaster management, to name just a few.

5G and beyond networks: The 5G has been promised for more bandwidth and high speed to accommodate the growing demands for public desire to connect to anything anywhere and at any time. Such stimulations impose a great pressure on the frequency spectrum as it is currently heavily consumed and no other alternatives seem feasible except transfer to new, higher bands. Such concept is also principally applicable to the WiMAX systems, as both are migrating to higher frequency bands that were not explored before. Whether in the new or old bands, the CR systems and CRNs have always been seen as very promising candidates to render more efficient and reliable spectrum use, access and management. Also, new technologies are nowadays trending, such as massive MIMO, which are worth further study in this regard.

Smart transportation: There has been a growing demand for efficient strategies concerning bus rapid transit, light rail transit including computer-based train control, vehicle-to-vehicle, and vehicle-to-infrastructure witnessed among others lately. This will naturally put high pressure on the spectrum bands that are currently operational or to come up with new bands to satisfy such new wireless platforms. The CR systems and CRNs have been proved to be of great potentials to support such modernised applications and thereby deserve more attention for focused elaborations.

Smart power and grids: The power, grid and utility networks are worth for better organisation, security and centralisation of operation, maintenance and upgradability options. On many scales and occasions, the CR systems and CRNs have well been treated in favour of such applications and the expansion of which is expected to be explosive soon. The power grids, whether of generation, transmission or distribution nature, have always sought to integrate with more autonomous and smart options to save time and resources in the multifaceted operations and services.

References

[1] S. Haykin, “Cognitive Radio: Brain-Empowered Wireless Communications”, *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 2, 2005, p. 201-220.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran and S. Mohanty, “NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey”, *Computer Networks*, Vol. 50, 2006, p. 2127-2159.

[3] A. Martian, C. Vladeanu, I. Marcu and I. Marghescu, “Evaluation of Spectrum Occupancy in an Urban Environment in a Cognitive Radio Context”, *International Journal on Advances in Telecommunications*, Vol. 3 no 3 & 4, 2010, p. 172-181.

[4] M. Fitch, M. Nekovee, S. Kawade, K. Briggs and R. MacKenzie, “Wireless Service Provision in TV White Space with Cognitive Radio Technology: A Telecom Operator’s Perspective and Experience”, *IEEE Communications Magazine*, Mar. 2011, p. 64-73.

[5] M. Nekovee, T. Irnich and R. Karlsson, “Worldwide Trends in Regulation of Secondary Access to White Spaces Using Cognitive Radio”, *IEEE Communications Magazine*, Aug. 2012, p. 32-40.

[6] C. Ghosh, S. Roy and D. Cavalcanti, “Coexistence Challenges for Heterogeneous Cognitive Wireless Networks in TV White Spaces”, *IEEE Wireless Communications*, Aug. 2011, p. 22-31.

[7] T. Baykas, M. Kasslin, M. Cummings, H. Kang, J. Kwak, R. Paine, A. Reznik, R. Saeed and S. J. Shellhammer, “Developing a Standard for TV White Space Coexistence: Technical Challenges and Solution Approaches”, *IEEE Wireless Communications*, Feb. 2012, p. 10-22.

[8] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shellhammer and W. Caldwell, “IEEE 802.22: The First Cognitive Radio Wireless Regional Area Network Standard”, *IEEE Communications Magazine*, Jan. 2009, p. 130-138.

[9] J. J. Popoola and R. van Olst, “A Novel Modulation-Sensing Method”, *IEEE Vehicular Technology Magazine*, Sep. 2011, p. 60-69.

[10] F. K. Jondral, “Cognitive Radio: A Communications Engineering View”, *IEEE Communications Magazine*, Aug. 2007, p. 28-33.

[11] R. Baldemair, E. Dahlman, G. Fodor, G. Mildh, S. Parkvall, Y. Selen, H. Tullberg and K. Balachandran, “Evolving Wireless Communications: Addressing the Challenges and Expectations of the Future”, *IEEE Vehicular Technology Magazine*, Mar. 2013, p. 24-30.

[12] L. Hanzo, H. Haas, S. Imre, D. O’Brien, M. Rupp and L. Gyongyosi, “Wireless Myths, Realities, and Futures: From 3G/4G to Optical and Quantum Wireless”, *Proceedings of the IEEE*, Vol. 100, May 2012, p. 1853-1888.

[13] D. Raychaudhuri and N. B. Mandayam, “Frontiers of Wireless and Mobile Communications”, *Proceedings of the IEEE*, Vol. 100, No. 4, Apr. 2012, p. 824-840.

[14] R. Venkatesha Prasad, P. Pawelczak, J. A. Hoffmeyer and H. Steven Berger, “Cognitive Functionality in Next Generation Wireless Networks: Standardization Efforts”, *IEEE Communications Magazine*, Apr. 2008, p. 72-78.

[15] B. Fette, “*Cognitive Radio Technology*”, Academic Press, 2009.

[16] H. Arslan (Ed.), “*Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems*”, Springer, 2007.

[17] C. Cordeiro, K. Challapali, D. Birru and S. Shankar N, “IEEE 802.22: The First Worldwide Wireless Standard based on Cognitive Radios”, *Proceedings of DySPAN*, 8-11 Nov. 2005, Baltimore, USA.

[18] M. Sherman, A. N. Mody, R. Martinez, C. Rodriguez and R. Reddy, “IEEE Standards Supporting Cognitive Radio and Networks, Dynamic Spectrum Access, and Coexistence”, *IEEE Communications Magazine*, Jul. 2008, p. 72-79.

[19] P. Gronsund, P. Pawelczak, J. Park and D. Cabric, “System Level Performance of IEEE 802.22-2011 with Sensing-Based Detection of Wireless Microphones”, *IEEE Communications Magazine*, Jan. 2014, p. 200-209.

[20] V. Stavroulaki, K. Tsagkaris, P. Demestichas, J. Gebert, M. Mueck, A. Schmidt, R. Ferrus and O. Sallent, “Cognitive Control Channels: From Concept to Identification of Implementation Options”, *IEEE Communications Magazine*, Jul. 2012, p. 96-108.

[21] W. Bolton, Y. Xiao and M. Guizani, “IEEE 802.20: Mobile Broadband Wireless Access”, *IEEE Wireless Communications*, Feb. 2007, p. 84-95.

[22] A. Greenspan, M. Klerer, J. Tomcik, R. Canchi and J. Wilson, “IEEE 802.20: Mobile Broadband Wireless Access for the Twenty-First Century*”, IEEE Communications Magazine*, Jul. 2008, p. 56-63.

[23] A. D. Olica, A. Banchs, I. Soto, T. Melia and A. Vidal, “An Overview of IEEE 802.21: Media Independent Handover Services”, *IEEE Wireless Communications*, Aug. 2008, p. 96-103.

[24] J. Mitola, “The Software Radio Architecture”, *IEEE Communications Magazine*, May 1995, p. 26-38.

[25] J. Mitola and G. Q. Maguire, “Cognitive Radio: Making Software Radios More Personal”, *IEEE Personal Communications*, Aug. 1999, p. 13-18.

[26] J. Mitola, “Software Radio Architecture: A Mathematical Perspective”, *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 4, Apr. 1999, p. 514-538.

[27] J. Mitola, “Software Radio Architecture Evolutions: Foundations, Technology Tradoffs, and Architecture Implications”, *IEICE Transactions on Communications*, Vol. E83-B, No. 6, Jun. 2000, p. 1165-1173.

[28] J. Mitola, “Cognitive Radio for Flexible Mobile Multimedia Communications”, *Mobile Networks and Applications*, Vol. 6, 2001, p. 435-441.

[29] S. Srikanteswara, R. C. Palat, J. H. Reed and P. Athanas, “An Overview of Configurable Computing Machines for Software Radio Handsets”, *IEEE Communications Magazine*, Jul. 2003, p. 134-141.

[30] M. L. Dickens, B. P. Dunn and J. N. Laneman, “Design and Implementation of a Portable Software Radio”, *IEEE Communications Magazine*, Aug. 2008, p. 58-66.

[31] A. B. MacKenzie, J. H. Reed, P. Athanas, C. W. Bostian, R. M. Buehrer, L. A. DaSilva, S. W. Ellingson, Y. T. Hou, M. Hsiao, J.-M. Park, C. Patterson, S. Raman and C. R. C. M. da Silva, "Cognitive Radio and Networking Research at Virginia Tech", *Proceedings of the IEEE*, Vol. 97, No. 4, Apr. 2009, p. 660-688.

[32] A. A. Abidi, "The Path to the Software-Defined Radio Receiver", *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 5, May 2007, p. 954-966.

[33] G. Hueber, Y. Zou, K. Dufrene, R. Stuhlberger and M. Valkama, "Smart Front-End Signal Processing for Advanced Wireless Receivers", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 3, No. 3, Jun. 2009, p. 472-487.

[34] Y. Zhao, S. Mao, J. O. Neel and J. H. Reed, "Performance Evaluation of Cognitive Radios: Metrics, Utility Functions, and Methodology", *Proceedings of the IEEE*, Vol. 97, No. 4, Apr. 2009, p. 642-659.

[35] J. Mitola, "Cognitive Radio Architecture Evolution", *Proceedings of the IEEE*, Vol. 97, No. 4, Apr. 2009, p. 626-641.

[36] Y. Tachwali, F. Basma and H. H. Refai, "Cognitive Radio Architecture for Rapidly Deployable Heterogeneous Wireless Networks", *IEEE Transactions on Consumer Electronics*, Vol. 56, No. 3, August 2010, p. 1426-1432.

[37] S. Srikanteswara, J. H. Reed, P. Athanas and R. Boyle, "A Soft Radio Architecture for Reconfigurable Platforms", *IEEE Communications Magazine*, Feb. 2000, p. 140-147.

[38] D. Makris, G. Gardikis and A. Kourtis, "Quantifying TV White Space Capacity: A Geolocation-Based Approach", *IEEE Communications Magazine*, Sep. 2012, p. 145-152.

[39] A. Goldsmith, "*Wireless Communications*", Cambridge, 2005.

[40] L. Hanzo, C. H. Wong and M. S. Yee, "*Adaptive Wireless Transceivers*", John Wiley & Sons Ltd, 2002.

[41] A. Zalonis, N. Miliou, I. Dagres, A. Polydoros and H. Bogucka, "Trends in Adaptive Modulation and Coding", *Advances in Electronics and Telecommunications*, Vol. 1, No. 1, Apr. 2010, p. 104-111.

[42] E. Biglieri and P. Di Torino, “Digital Transmission in 21st Century: Conflating Modulation and Coding”, *IEEE Communication Magazine*, May 2002, p. 128-137.

[43] A. Svensson, “An Introduction to Adaptive QAM Modulation Schemes for Known and Predicted Channels”, *Proceedings of the IEEE*, Vol. 95, No. 12, Dec. 2007, p. 2322-2336.

[44] G. Caire and K. R. Kumar, “Information Theoretic Foundations of Adaptive Coded Modulation”, *Proceedings of the IEEE*, Vol. 95, No. 12, Dec. 2007, p. 2274-2298.

[45] E. Biglieri, J. Proakis and S. Shamai, “Fading Channels: Information-Theoretic and Communications Aspects”, *IEEE Transactions on Information Theory*, Vol. 44, No. 6, Oct. 1998, p. 2619-2692.

[46] M. K. Ozdemir and H. Arslan, “Channel Estimation for Wireless OFDM Systems”, *IEEE Communications Surveys & Tutorials*, 2nd Q. 2007, p. 18-48.

[47] A. J. Goldsmith and S.-G. Chua, “Variable-Rate Variable-Power MQAM for Fading Channels”, *IEEE Transactions on Communications*, Vol. 45, No. 10, Oct. 1997, p. 1218-1230.

[48] S. Falahati, A. Svensson, T. Ekman and M. Sternad, “Adaptive Modulation Systems for Predicted Wireless Channels”, *IEEE Transactions on Communications*, Vol. 52, No. 2, Feb. 2004, p. 307-316.

[49] A. J. Goldsmith and S.-G. Chua, “Adaptive Coded Modulation for Fading Channels”, *IEEE Transactions on Communications*, Vol. 46, No. 5, May 1998, p. 595-602.

[50] S. Falahati, A. Svensson, M. Sternad and H. Mei, “Adaptive Trellis-Coded Modulation over Predicted Flat Fading Channels”, *Proc. IEEE VTC*, Orlando, Oct. 2003, p. 1532-1536.

[51] S. T. Chung and A. J. Goldsmith, “Degrees of Freedom in Adaptive Modulation: A Unified View”, *IEEE Transactions on Communications*, Vol. 49. No. 9, Sep. 2001, p. 1561-1571.

[52] P. Ormeci, X. Liu, D. L. Goeckel and R. D. Wesel, “Adaptive Bit-Interleaved Coded Modulation”, *IEEE Transactions on Communication Systems*, Vol. 49, No. 9, Sep. 2001, p. 1572-1581.

[53] S. Vishwanath and A. Goldsmith, “Adaptive Turbo-Coded Modulation for Flat-Fading Channels”, *IEEE Transactions on Communications*, Vol. 51, No. 6, Jun. 2003, p. 964-972.

[54] R. Sassioui, M. Jabi, L. Szczecinski, L. B. Le, M. Benjillali and B. Pelletier, “HARQ and AMC: Friends or Foes?”, *IEEE Transactions on Communications*, Vol. 65, No. 2, Feb. 2017, p. 635-650.

[55] Y.-C. Liang, K.-C. Chen, G. Y. Li and P. Mahonen, “Cognitive Radio Networking and Communications: An Overview”, *IEEE Transactions on Vehicular Technology*, Vol. 60, No. 7, Sep. 2011, p. 3386-3407.

[56] A. Tsakmalis, S. Chatzinotas and B. Ottersten, “Centralized Power Control in Cognitive Radio Networks Using Modulation and Coding Classification Feedback”, *IEEE Transactions on Cognitive Communications and Networking*, Vol. 2, No. 3, Sep. 2016, p. 223-237.

[57] W. Liang, S. X. Ng and L. Hanzo, “Cooperative Communication between Cognitive and Primary Users”, *IET Communications*, Vol. 7, No. 17, 2013, p. 1982–1992.

[58] W. Liang, S. X. Ng and L. Hanzo, “Cooperative Overlay Spectrum Access in Cognitive Radio Networks”, *IEEE Communications Surveys & Tutorials*, Vol. 19, No. 3, 3rd Q. 2017, p. 1924-1944.

[59] O. A. Dobre, A. Abdi, Y. B.-Ness and W. Su, “Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends”, *IET Communications*, Vol. 1, No. 2, Apr. 2007, p. 137-156.

[60] Y. A. Eldemerdash, O. A. Dobre and M. Oner, “Signal Identification for Multiple-Antenna Wireless Systems: Achievements and Challenges”, *IEEE Communications Surveys & Tutorials*, Vol. 18, No. 3, 3rd Q. 2016, p. 1524-1551.

[61] O. A. Dobre, “Signal Identification for Emerging Intelligent Radios: Classical Problems and New Challenges”, *IEEE Instrumentation & Measurement Magazine*, Apr. 2015, p. 11-18.

[62] M. R. Bahloul, M. Z. Yusoff, A.-H. A.-Aty, M. N. M. Saad and M. Al-Jemeli, “Modulation Classification for MIMO Systems: State of the Art and Research Directions”, *Chaos, Solitons and Fractals*, Vol. 89, 2016, p. 497–505.

[63] B. Ramkumar, “Automatic Modulation Classification for Cognitive Radio Using Cyclic Feature Detection”, *IEEE Circuits and Systems*, 2nd Q. 2009, p. 27-45.

[64] Z. Zhu and A. K. Nandi, “*Automatic Modulation Classification: Principles, Algorithms, and Applications*”, John Wile & Sons, 2015.

[65] B. Li, S. Li, J. Hou, J. Fu, C. Zhao and A. Nallanathan, “A Bayesian Approach for Adaptively Modulated Signals Recognition in Next-Generation Communications”, *IEEE Transactions on Signal Processing*, Vol. 63, No. 16, Aug. 2015, p. 4359-4372.

[66] A. Gorcin and H. Arslan, “Signal Identification for Adaptive Spectrum Hyperspace Access in Wireless Communications Systems”, *IEEE Communications Magazine*, Oct. 2014, p. 134-145.

[67] M. Bkassiny, S. K. Jayaweera, Y. Li and K. A. Avery, “Wideband Spectrum Sensing and Non-Parametric Signal Classification for Autonomous Self-Learning Cognitive Radios”, *IEEE Transactions on Wireless Communications*, Vol. 11, No. 7, Jul. 2012, p. 2596-2605.

[68] O. A. Dobre, A. Abdi, Y. B.-Ness and W. Su, “Blind Modulation Classification: A Concept Whose Time Has Come”, *Proc. IEEE AWWC*, 18-19 Apr. 2005, p. 1-6.

[69] M. Marey and O. A. Dobre, “Blind Modulation Classification Algorithm for Single and Multiple-Antenna Systems Over Frequency-Selective Channels”, *IEEE Signal Processing Letters*, Vol. 21, No. 9, Sep. 2014, p. 1098-1102.

[70] S. Kharbech, I. Dayoub, M. Z.-Colin and E. P. Simon, “On Classifiers for Blind Feature-Based Automatic Modulation Classification Over Multiple-Input

Multiple-Output Channels", *IET Communications*, Vol. 10, Iss. 7, 2016, p. 790-795.

[71] G. Jajoo, Y. Kumar, S. K. Yadav, B. Adhikari and A. Kumar, "Blind Signal Modulation Recognition Through Clustering Analysis of Constellation Signature", *Expert Systems With Applications*, Vol. 90, 2017, p. 13-22.

[72] Z. Zhu and A. K. Nandi, "Blind Digital Modulation Classification Using Minimum Distance Centroid Estimator and Non-Parametric Likelihood Function", *IEEE Transactions on Wireless Communications*, Vol. 13, No. 8, Aug. 2014, p. 4483-4494.

[73] W. Su, J. L. Xu and M. Zhou, "Real-time Modulation Classification Based On Maximum Likelihood", *IEEE Communications Letters*, Vol. 12, No. 11, Nov. 2008, p. 801-803.

[74] F. Hameed, O. A. Dobre and D. C. Popescu, "On the Likelihood-Based Approach to Modulation Classification", *IEEE Transactions on Wireless Communications*, Vol. 8, No. 12, Dec. 2009, p. 5884-5892.

[75] J. L. Xu, W. Su and M. Zhou, "Software-Defined Radio Equipped With Rapid Modulation Recognition", *IEEE Transactions on Vehicular Technology*, Vol. 59, No. 4, May 2010, p. 1659-1667.

[76] J. L. Xu and M. Zhou, "Likelihood-Ratio Approaches to Automatic Modulation Classification", *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, Vol. 41, No. 4, Jul. 2011, p. 455-469.

[77] V. G. Chavali and R. C. M. da Silva, "Maximum-Likelihood Classification of Digital Amplitude-Phase Modulated Signals in Flat Fading Non-Gaussian Channels", *IEEE Transactions on Communications*, Vol. 59, No. 8, Aug. 2011, p. 2051-2056.

[78] M. Derakhtian, A. A. Tadaion and S. Gazor, "Modulation Classification of Linearly Modulated Signals in Slow Flat Fading Channels", *IET Signal Processing*, 2011, Vol. 5, Iss. 5, p. 443-450.

[79] A. Abdelmutalab, K. Assaleh and M. El-Tarhuni, “Automatic Modulation Classification Based on High Order Cumulants and Hierarchical Polynomial Classifiers”, *Physical Communication*, Vol. 21, 2016, p. 10-18.

[80] F. Wang and X. Wang, “Fast and Robust Modulation Classification via Kolmogorov-Smirnov Test”, *IEEE Transactions on Communications*, Vol. 58, No. 8, Aug. 2010, p. 2324-2332.

[81] S. A. Ghauri, I. M. Qureshi and A. N. Malik, “A Novel Approach for Automatic Modulation Classification via Hidden Markov Models and Gabor Features”, *Wireless Personal Communications*, May 2017, p. 1-18.

[82] S. Yan, Q. Song, J. Fu and C. Zhao, “Particle Filter Based Bayesian Inference Modulation Recognition Algorithm”, *Proc. ISCIT*, 26-28 Sep. 2016, Qingdao, China, p. 96-99.

[83] Y. Liu, O. Simeone, A. M. Haimovich and W. Su, “Modulation Classification via Gibbs Sampling Based on a Latent Dirichlet Bayesian Network”, *IEEE Signal Processing Letters*, Vol. 21, No. 9, Sep. 2014, p. 1135-1139.

[84] S. Norouzi, A. Jamshidi and A. R. Zolghadrasli, “Adaptive Modulation Recognition Based on the Evolutionary Algorithms”, *Applied Soft Computing*, Vol. 43, 2016, p. 312-319.

[85] M. Aladeemy, S. Tutun and M. T. Khasawneh, “A New Hybrid Approach for Feature Selection and Support Vector Machine Model Selection Based on Self-Adaptive Cohort Intelligence”, *Expert Systems with Applications*, Vol. 88, 2017, p. 118-131.

[86] J. L. Xu, W. Su and M. Zhou, “Distributed Automatic Modulation Classification With Multiple Sensors”, *IEEE Sensors Journal*, Vol. 10, No. 11, Nov. 2010, p. 1779-1785.

[87] W. Su, “Modulation Classification of Single-Input Multiple-Output Signals Using Asynchronous Sensors”, *IEEE Sensors Journal*, Vol. 15, No. 1, Jan. 2015, p. 346-357.

[88] B. Dulek, “An Online and Distributed Approach for Modulation Classification Using Wireless Sensor Networks”, *IEEE Sensors Journal*, Vol. 17, No. 6, Mar, 2017, p. 1781-1787.

[89] A. Ghasemi and E. S. Sousa, “Spectrum Sensing in Cognitive Radio Networks: Requirements, Challenges and Design Trade-offs”, *IEEE Communications Magazine*, Apr. 2008, p. 32-39.

[90] T. Yucek and H. Arslan, “A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications”, *IEEE Communications Surveys and Tutorials*, Vol. 11, No. 1, 1st Quarter 2009, pp. 116-130.

[91] B. Wang and K. Liu, “Advances in Cognitive Radio Networks: A Survey*”, IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No. 1, Feb. 2011, pp. 5-23.

[92] E. Axell, G. Leus, E. G. Larsson and H. V. Poor, “Spectrum Sensing for Cognitive Radio: State of the Art and Recent Advances”, *IEEE Signal Processing Magazine*, May 2012, pp. 101-116.

[93] J. Ma, G. Y. Li and B. H. Juang, “Signal Processing in Cognitive Radio”, *Proceedings of the IEEE*, Vol. 97, No. 5, May 2009, pp. 805-823.

[94] E. Hossain, D. Niyato and D. I. Kim, “Evolution and Future Trends of Research in Cognitive Radio: A Contemporary Survey”, *Wireless Communications and Mobile Computing*, Vol. 15, 2015, pp. 1530-1564.

[95] R. Umara and A. U. H. Sheikh, “A Comparative Study of Spectrum Awareness Techniques for Cognitive Radio Oriented Wireless Networks”, *Physical Communication*, Vol. 9, 2013, pp. 148–170.

[96] Y. Zeng, Y.-C. Liang, A. T. Hoang and R. Zhang, “A Review on Spectrum Sensing for Cognitive Radio: Challenges and Solutions”, *EURASIP Journal on Advances in Signal Processing*, Vol. 2010, Article ID 381465, pp. 1-15.

[97] L. Claudino and T. Abrao, “Spectrum Sensing Methods for Cognitive Radio Networks: A Review”, *Wireless Personal Communications*, Apr. 2017, p. 1-35.

[98] S. K. Sharma, T. E. Bogale, S. Chatzinotas, B. Ottersten, L. B. Le and X. Wang, "Cognitive Radio Techniques Under Practical Imperfections: A Survey", *IEEE Communication Surveys & Tutorials*, Vol. 17, No. 4, 4th Q. 2015, p. 1858-1884.

[99] W. Ejaz, N. Hasan, M. A. Azam and H. S. Kim, "Improved Local Spectrum Sensing for Cognitive Radio Networks", *EURASIP Journal on Advances in Signal Processing*, Vol. 242, 2012, p. 1-12.

[100] K. B. Letaief and W. Zhang, "Cooperative Communications for Cognitive Radio Networks", *Proceedings of the IEEE*, Vol. 97, No. 5, May 2009, p. 878-893.

[101] L. Gavrilovska and V. Atanasovski, "Spectrum Sensing Framework for Cognitive Radio Networks", *Wireless Personal Communications*, 2011, Vol. 59, p. 447–469.

[102] I. F. Akyildiz, B. F. Lo and R. Balakrishnan, "Cooperative Spectrum Sensing in Cognitive Radio Networks: A Survey", *Physical Communication*, Vol. 4, 2011, p. 40-62.

[103] E. Hossain, D. Niyato and Z. Han, "*Dynamic Spectrum Access and Management in Cognitive Radio Networks*", Cambridge, 2009.

[104] M. Ibnkahla, "*Cooperative Cognitive Radio Networks: The Complete Spectrum Cycle*", CRC Press, 2015.

[105] S. Atapattu, C. Tellambura and H. Jiang, "*Energy Detection for Spectrum Sensing in Cognitive Radio*", Springer, 2014.

[106] S. Haykin, "*Cognitive Dynamic Systems: Perception-Action, Cycle, Radar, and Radio*", Cambridge University Press, 2012.

[107] J. G. Proakis and M. Salehi, "*Digital Communications*", McGraw-Hill, 2001.

[108] D. G. Brennan, "Linear Diversity Combining Techniques", *Proceedings of the IEEE,* Vol. 91, No. 2, Feb. 2003, p. 331-356..

[109] C. R. C. M. da Silva and M. D. Yacoub, "A Generalized Solution for Diversity Combining Techniques in Fading Channels", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, No. 1, Jan. 2002, p. 46-50.

[110] T. K. Y. Lo, "Maximum Ratio Transmission", *IEEE Transactions on Communications*, Vol. 47, No. 10, Oct. 1999, p. 1458-1461.

[111] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker and P. A. Hoeher, "Multiple-Antenna Techniques for Wireless Communications - A Comprehensive Literature Survey", *IEEE Communications Surveys & Tutorials*, Vol. 11, No. 2, 2nd Q. 2009, p. 87-105.

[112] T. Hwang, C. Yang, G. Wu, S. Li and G. Y. Li, "OFDM and Its Wireless Applications: A Survey", *IEEE Transactions on Vehicular Technology*, Vol. 58, No. 4, May 2009, p. 1673-1694.

[113] G. L. Stuber, J. R. Barry, S. W. McLaughlin, Y. LI, M. A. Ingram and T. G. Pratt, "Broadband MIMO-OFDM Wireless Communications", *Proceedings of the IEEE*, Vol. 92, No. 2, Feb. 2004, p. 271-294.

[114] A. El-Mougy, M. Ibnkahla, G. Hattab and W. Ejaz, "Reconfigurable Wireless Networks", *IEEE of the Proceedings*, Vol. 103, No. 7, Jul. 2015, p.1125-1158.

[115] A. Naeem, M. H. Rehmani, Y. Saleem, I. Rashid and N. Crespi, "Network Coding in Cognitive Radio Networks: A Comprehensive Survey", *IEEE Communications Surveys & Tutorials*, Vol. 19, No. 3, 3rd Q., 2017, p. 1945-1973.

[116] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications", *IEEE Journal on Selected Areas in Communications,* Vol. 16, No. 8, Oct. 1998, p. 1451–1458.

[117] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction", *IEEE Transactions on Information Theory*, Vol. 44, No. 2, Mar 1998, p. 744–765.

[118] G. B. Giannakis, Z. Liu, X. Ma and S. Zhou, "*Space-Time Coding for Broadband Wireless Communications*", John Wiley & Sons, 2007.

[119] A. Paulraj, R. Nabar and D. Gore, "*Introduction to Space-Time Wireless Communications*", Cambridge University Press, 2003.

[120] H. Jafarkhani, "Quasi-Orthogonal Space-Time Block Code", *IEEE Transactions on Communications*, Vol. 49, No. 1, Jan. 2001, p. 1-4.

[121] B. F.-Boroujeny and R. Kempter, “Multicarrier Communication Techniques for Spectrum Sensing and Communication in Cognitive Radios”, *IEEE Communications Magazine*, Apr. 2008, p. 80-85.

[122] I. Budiarjo, H. Nikookar and L. P. Ligthart, “Cognitive Radio Modulation Techniques”, *IEEE Signal Processing Magazine*, Nov. 2008, p. 24-34.

[123] H. A. Mahmoud, T. Yucek and H. Arslan, “OFDM for Cognitive Radio: Merits and Challenges”, *IEEE Wireless Communications*, Apr. 2009, p. 6-14.

[124] H. Bolcskei and E. Zurich, “MIMO-OFDM Wireless Systems: Basics, Perspectives, and Challenges”, *IEEE Wireless Communications*, Aug. 2006, p. 31-37.

[125] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, “Finite-State Markov Modeling of Fading Channels: A Survey of Principles and Applications”, *IEEE Signal Processing Magazine*, Sep. 2008, p. 57-80.

[126] H. S. Wang and N. Moayeri, “Finite-State Markov Channel-A Useful Model for Radio Communication Channels”, *IEEE Transactions on Vehicular Technology*, Vol. 44, No. 1, Feb. 1995, p. 163-171.

[127] Q. Zhang and S. A. Kassam, “Finite-State Markov Model for Rayleigh Fading Channels”, *IEEE Transactions on Communications*, Vol. 47, No. 11, Nov. 1999, p. 1688-1692.

[128] F. Babich, O. E. Kelly, and G. Lombardi, “Generalised FSMC Model for Radio Channels with Correlated Fading”, *IEEE Transactions on Communications*, Vol. 48, No. 4, Apr. 2000, p. 547-551.

[129] C. C. Tan and N. C. Beaulieu, “On First-Order Markov Modeling for the Rayleigh Fading Channel”, *IEEE Transactions on Communications*, Vol. 48, No. 12, Dec. 2000, p. 2032-2040

[130] J. M. Park and Gang U. Hwang, “Mathematical Modeling of Rayleigh Fading Channels Based on Finite State Markov Chains”, *IEEE Communications Letters*, Vol. 13, No. 10, Oct. 2009, p. 764-766.

[131] D. Dalalah, L. Cheng and G. Tonkay, "Modeling End-to-End Wireless Lossy Channels: A Finite-State Markov Approach", *IEEE Transactions on Wireless Communications*, Vol 7, No. 4, Apr. 2008, p. 1236-1243.

[132] S. F.-Dehkordy, J. Abouei, K. Plataniotis and S. Pasupathy, "Markovian-Based Framework for Cooperative Channel Selection in Cognitive Radio Networks*",* *IET Communications*, 2014, Vol. 8, Iss. 14, p. 2458–2468.

[133] K. E. Baddour and N. C. Beaulieu, "Autoregressive Modeling for Fading Channel Simulation", *IEEE Transactions on Wireless Communications*, Vol. 4, No. 4, Jul. 2005, p. 1650-1662.

[134] H. Mehrpouyan and S. D. Blostein, "ARMA Synthesis of Fading Channels", *IEEE Transactions on Wireless Communications*, Vol. 7, No. 8, Aug. 2008, p. 2846-2850.

[135] G. Tauchen, "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions", *Economic Letters*, Vol. 20, 1986, p. 177-181.

[136] K. A. Kopecky and R. M. H. Suen, "Finite State Markov-Chain Approximations to Highly Persistent Processes", *Review of Economic Dynamics*, Vol. 13, 2010, p. 701–714.

[137] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proceedings of the IEEE*, Vol. 77, No. 2, Feb. 1989, p. 257-286.

[138] W. Khreich, E. Granger, A. Miri and R. Sabourin, "On the Memory Complexity of the Forward-Backward Algorithm", *Pattern Recognition Letters*, Vol. 31, 2010, p. 91–99.

[139] F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark, "*NIST Handbook of Mathematical Functions*", NIST and Cambridge University Press, 2010.

[140] F. Khozeimeh and S. Haykin, "Brain-Inspired Dynamic Spectrum Management for Cognitive Radio Ad Hoc Networks", *IEEE Transactions on Wireless Communications*, Vol. 11, No. 10, Oct. 2012, p. 3509-3517.

[141] S. Haykin and P. Setoodeh, "Cognitive Radio Networks: The Spectrum Supply Chain Paradigm", *IEEE Transactions on Cognitive Communications and Networking*, Vol. 1, No. 1, Mar. 2015, p. 3-28.

[142] L. L. Scharf, "*Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*", Addison Wesley, 1991.

[143] S. M. Kay, "*Fundamentals of Statistical Signal Processing: Estimation Theory*", Vol. I, Prentice Hall, 1993.

[144] H. Arslan and G. E. Bottomley. "Channel Estimation in Narrowband Wireless Communication Systems", *Wireless Communications and Mobile Computing*, Vol. 1, 2001, p. 201-219.

[145] Y. Liu, Z. Tan, H. Hu, L. J. Cimini, Jr. and G. Y. Li, "Channel Estimation for OFDM", *IEEE Communications Surveys and Tutorials*, Vol. 16, No. 4, 4th Q. 2014, p. 1891-1908.

[146] C. Shin, R. W. Heath, Jr. and E. J. Powers, "Blind Channel Estimation for MIMO-OFDM Systems", *IEEE Transactions on Vehicular Technology*, Vol. 56, No. 2, Mar. 2007, p. 670-685.

[147] X. Ma, G. B. Giannakis and S. Ohno, "Optimal Training for Block Transmissions Over Doubly Selective Wireless Fading Channels", *IEEE Transactions on Signal Processing*, Vol. 51, No. 5, May 2003, p. 1351-1366.

[148] S. Haykin, K. Huber and Z. Chen, "Bayesian Sequential State Estimation for MIMO Wireless Communications", *Proceedings of the IEEE*, Vol. 92, No. 3, Mar. 2004, p. 439-454.

[149] C. Komninakis, C. Fragouli, A. H. Sayed and R. D. Wesel, "Multi-Input Multi-Output Fading Channel Tracking and Equalization Using Kalman Estimation", *IEEE Transactions on Signal Processing*, Vol. 50, No. 5, May 2002, p. 1065–1076.

[150] K.-Y. Han, S.-W. Lee, J.-S. Lim and K.-M. Sung, "Channel Estimation for OFDM with Fast Fading Channels by Modified Kalman Filter", *IEEE Transactions on Consumer Electronics*, Vol. 50, No. 2, May 2004, p. 444-449.

[151] T. Y. Al-Naffouri, “An EM-Based Forward-Backward Kalman Filter for the Estimation of Time-Variant Channels in OFDM”, *IEEE Transactions on Signal Processing*, Vol. 55, No. 7, Jul. 2007, p. 3924-3930.

[152] J. Choi, A. C. de C. Lima and S. Haykin, “Kalman Filter-Trained Recurrent Neural Equalizers for Time-Varying Channels”, *IEEE Transactions on Communications*, Vol. 53, No. 3, Mar. 2005, p. 472-480.

[153] S. Noh, Michael D. Zoltowski and D. J. Love, “Training Sequence Design for Feedback Assisted Hybrid Beamforming in Massive MIMO Systems”, *IEEE Transactions on Communications*, Vol. 64, No. 1, Jan. 2016, p. 187-200.

[154] A. Ribeiro, I. D. Schizas, S. I. Roumeliotis and G. B. Giannakis, “Kalman Filtering in Wireless Sensor Networks”, *IEEE Control Systems Magazine*, Apr. 2010, p. 66-86.

[155] L. A. Johnston and V. Krishnamurthy, “An Improvement to the Interacting Multiple Model (IMM) Algorithm”, *IEEE Transactions on Signal Processing*, Vol. 49, No. 12, Dec. 2001, p. 2909-2923.

[156] T. Sathyan and T. Kirubarajan, “Markov-Jump-System-Based Secure Chaotic Communication”, *IEEE Transactions on Circuits and Systems-I: Regular Paper*, Vol. 53, No. 7, Jul. 2006, p. 1597-1609.

[157] E. Mazor, A. Averbuch, Y. B.-Shalom and J. Dayan, “Interacting Multiple Model Methods in Target Tracking: A Survey”, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 34, No. 1, Jan. 1998, p. 103-123.

[158] H. Zhang, S. Wei, G. Ananthaswamy and D. L. Goeckel, “Adaptive Signaling Based on Statistical Characterizations of Outdated Feedback in Wireless Communications”, *Proceedings of the IEEE*, Vol. 95, No. 12, Dec. 2007, p. 2337-2353.

[159] S. H. Ting, K. Sakaguchi and K. Araki, “A Robust and Low Complexity Adaptive Algorithm for MIMO Eigenmode Transmission System with Experimental Validation”, *IEEE Transactions on Wireless Communications*, Vol. 5, No. 7, Jul. 2006, p. 1775-1784.

[160] M. Z. Shakir and T. S. Durrani, “MIMO Identical Eigenmode Transmission (IETS) – A Channel Decomposition Perspective”, *Proc. EUSIPCO*, 3-7 Sep. 2007,Poznan, p. 916-920.

[161] L. G. Ordonez, D. P. Palomar and J. R. Fonollosa, “Ordered Eigenvalues of a General Class of Hermitian Random Matrices and Performance Analysis of MIMO Systems”, *Proc. ICC*, 18-23 May 2008, Beijing, p. 3846-3852.

[162] L. G. Ordonez, D. P. Palomar, A. P.-Zamora and J. R. Fonollosa, “High-SNR Analytical Performance of Spatial Multiplexing MIMO Systems With CSI”, *IEEE Transactions on Signal Processing*, Vol. 55, No. 11, Nov. 2007, p. 5447-5463.

[163] S. Jin, M. R. McKay, X. Gao and I. B. Collings, “MIMO Multichannel Beamforming: SER and Outage Using New Eigenvalue Distributions of Complex Noncentral Wishart Matrices”, *IEEE Transactions on Communications*, Vol. 56, No. 3, Mar. 2008, p. 424-434.

[164] K. P. Kongara and P. J. Smith, “Performance Analysis of MIMO-SVD and MMSE-Receivers for Adaptive OFDM Systems”, *Proc. VTC*, 2-5 Sep. 2013, Las Vegas, p. 1-5.

[165] X. Yu and Y. Jing, “SVD-Based Channel Estimation for MIMO Relay Networks”, *Proc. VTC*, 3-6 Sep. 2012, Quebec, p. 1-5.

[166] X. Wang and S. Brink, “Iterative MIMO Subspace Detection based on Parallel Interference Cancellation”, *Proc. WCNC*, 19-20 Mar. 2017, San Francesco, p. 1-6.

[167] X. Wei, W. Peng, D. Chen, R. Schober and T. Jiang, “Uplink Channel Estimation in Massive MIMO Systems Using Factor Analysis”, *EA, IEEE Communications Letters*, May 2018, p. 1-4.

[168] S. S. Khalid and S. Abrar, “A Low-Complexity Interacting Multiple Model Filter for Maneuvering Target Tracking”, *International Journal of Electronics and Communications (AEÜ)*, Vol. 73, 2017, p. 157-164.

[169] E. Biglieri and P. Di Torino, “Coding and Modulation for a Horrible Channel”, *IEEE Communications Magazine*, May 2003, p. 92-98.

[170] G. Ungerboeck, “Channel Coding with Multilevel/Phase Signals”, *IEEE Transactions on Information Theory*, Vol. IT-28, No. 1, Jan. 1982, p. 55-67.

[171] G. D. Forney, Jr. and G. Ungerboeck, “Modulation and Coding for Linear Gaussian Channels”, IEEE Transactions on Information Theory, Vol. 44, No. 6, Oct. 1998, p. 2384-2415.

[172] E. Zehavi, “8-PSK Trellis Codes for a Rayleigh Channel”, *IEEE Transactions on Communications*, Vol. 40, No. 3, May 1992, p. 873-884.

[173] G. Caire, G. Taricco and E. Biglieri, “Bit-Interleaved Coded Modulation”, *IEEE Transactions on Information Theory*, Vol. 44, No. 3, May 1998, p. 927-946.

[174] A. Fabregas and G. Caire, “Coded Modulation in the Block-Fading Channel: Coding Theorems and Code Construction”, *IEEE Transactions on Information Theory*, Vol. 52, No. 1, Jan. 2006, p. 91-114.

[175] X. Li and J. A. Ritcey, “Bit-Interleaved Coded Modulation With Iterative Decoding,” *IEEE Communications Letters*, Vol. 1, No. 6, Nov. 1997, p. 169-171.

[176] X. Li, A. Chindapol and J. A. Ritcey, “Bit-Interleaved Coded Modulation With Iterative Decoding and 8PSK Signaling,” *IEEE Transactions on Communications*, Vol. 50, No. 6, Aug. 2002, p. 1250–1257.

[177] B. K. Khoo, S. Y. Le Goff, B. S. Sharif and C. C. Tsimenidis, “Bit-Interleaved Coded Modulation With Iterative Decoding Using Constellation Shaping”, *IEEE Transactions on Communications*, Vol. 54, No. 9, Sep. 2006, p. 1517-1520.

[178] D. L. Goeckel, “Adaptive Coding for Time-Varying Channels Using Outdated Fading Estimates”, *IEEE Transactions on Communications*, Vol. 47, No. 6, Jun. 1999, p. 844-855.

[179] V. K. N. Lau and M. D. Macleod, “Variable-Rate Adaptive Trellis Coded QAM for Flat-Fading Channels”, *IEEE Transactions on Communications*, Vol. 49, No. 9, Sep. 2001, p. 1550-1560.

[180] V. K. N. Lau, “Performance Analysis of Variable Rate: Symbol-by-Symbol Adaptive Bit Interleaved Coded Modulation for Rayleigh Fading Channels”, *IEEE Transactions on Vehicular Technology*, Vol. 51, No. 3, May 2002, p. 537-550.

[181] A. D.-Hallen, “Fading Channel Prediction for Mobile Radio Adaptive Transmission Systems”, *Proceedings of the IEEE*, Vol. 95, No. 12, Dec. 2007, p. 2299-2313.

[182] K. L. Baum, T. A. Kostas, P. J. Sartori and B. K. Classon, “Performance Characteristics of Cellular Systems With Different Link Adaptation Strategies”, *IEEE Transactions on Vehicular Technology*, Vol. 52, No. 6, Nov. 2003, p. 1479-1507.

[183] B. Vucetic, “An Adaptive Coding Scheme for Time-Varying Channels”, *IEEE Transactions on Communications*, Vol. 39, No. 5, May 1991, p. 653-663.

[184] T. L. Jensen, S. Kant, J. Wehinger and B. H. Fleury, “Fast Link Adaptation for MIMO OFDM”, *IEEE Transactions on Vehicular Technology*, Vol. 59, No. 8, Oct. 2010, p. 3766-3778.

[185] J. Chang, I.T. Lu1 and Y.X. Li, “Adaptive Codebook-Based Channel Prediction and Interpolation For Multiuser Multiple-Input Multiple-Output-Orthogonal Frequency Division Multiplexing Systems”, *IET Communications*, 2012, Vol. 6, Iss. 3, p. 281-288.

[186] U. F.-Plazaola, E. M.-Naya, J. F. Paris and A. J. Goldsmith, “Adaptive Modulation for MIMO Systems with Channel Prediction Errors”, *IEEE Transactions on Wireless Communications*, Vol. 9, No. 8, Aug. 2010, p. 2516-2527.

[187] A. Maaref and S Aissa, “Optimized Rate-Adaptive PSAM for MIMO MRC Systems with Transmit and Receive CSI Imperfections”, *IEEE Transactions on Communications*, Vol. 57, No. 3, Mar. 2009, p. 821-830.

[188] Z. Zhang, H. Luo, J. Zhang, W. Wang and G. Yu, “Cognitive Radio Transmission Strategies Exploiting the Primary-Link Adaptivity”, *IEEE Transactions on Vehicular Technology*, Vol. 60, No. 8, Oct. 2011, p. 3805-3813.

[189] J.-P. Hong and W. Choi, “Gains and Limits of Diversity Techniques in Cognitive Radio Systems”, *Journal of Communications and Networks*, Vol. 19, No. 2, Apr. 2017, p. 97-104.

[190] M. Torabi and C. Nerguizian, “Adaptive Transmission in Spectrum Sharing Systems With Alamouti OSTBC Under Spatially Correlated Channels”, *IEEE Transactions on Vehicular Technology*, Vol. 66, No. 4, Apr. 2017, p. 3131-3142.

[191] Y. Xu, X. Zhao and Y.-C. Liang, “Robust Power Control and Beamforming in Cognitive Radio Networks: A Survey”, *IEEE Communication Surveys & Tutorials*, Vol. 17, No. 4, 4th Q. 2015, p. 1834-1875.

[192] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar and A. Goldsmith, “Sum Power Iterative Water-Filling for Multi-Antenna Gaussian Broadcast Channels”, *IEEE Transactions on Information Theory*, Vol. 51, No. 4, Apr. 2005, p. 1570-1580.

[193] X. Li and J. A. Ritcey, “Trellis-Coded Modulation with Bit Interleaving and Iterative Decoding”, *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 4, Apr. 1999, p. 715-724.

[194] D. R. Pauluzzi and N. C. Beaulieu, “A Comparison of SNR Estimation Techniques for the AWGN Channel”, *IEEE Transactions on Communications*, Vol. 48, No. 10, Oct. 2000, p. 1681-1691.

[195] P. Gao and C. Tepedelenlioglu, “SNR Estimation for Nonconstant Modulus Constellations”, *IEEE Transactions on Signal Processing*, Vol. 53, No. 3, Mar. 2005, p. 865-870.

[196] A. Wiesel, J. Goldberg and H. M.-Yaron, “SNR Estimation in Time-Varying Fading Channels”, *IEEE Transactions on Communications*, Vol. 54, No. 5, May 2006, p. 841-848.

[197] R. B. Kerr, “On Signal and Noise Level Estimation in a Coherent PCM Channel,” *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-2, No. 3, Jul. 1966, p. 450-454.

[198] R. M. Gagliardi and C. M. Thomas, “PCM Data Reliability Monitoring Through Estimation of Signal-to-Noise Ratio,” *IEEE Transaction Communications*, Vol. COM-16, No. 6, Jun. 1968, p. 479-486.

[199] D. R. Joshi, D. C. Popescu and O. A. Dobre, “Gradient-Based Threshold Adaptation For Energy Detector in Cognitive Radio Systems’, *IEEE Communications Letters*, Vol. 15, No. 1, Jan. 2011, p. 19-21.

[200] A. Ijaz, A. B. Awoseyila and B. G. Evans, “Low-Complexity Time-Domain SNR Estimation For OFDM Systems”, *Electronics Letters*, Vol. 47, No. 20, Sep. 2011, p. 1-2.

[201] S. M. Kay, “*Fundamentals Statistical Signal Processing: Detection Theory*”, Vol. II, Prentice Hall-PTR, 1998.

[202] V. Choqueuse, S. Azou, K. Yao, L. Collin and G. Burel, “Blind Modulation Recognition For MIMO Systems,” *MTA Review*, Vol. 19, No. 2, Jun. 2009, p. 183-196.

[203] Z. Zhu and A. K. Nandi, “Blind Modulation Classification for MIMO systems using Expectation Maximization”, *Proc. MILCOM*, Baltimore, 6-8 Oct. 2014, p. 1-6.

[204] Q. Shi and Y. Karasawa, “Robust Maximum Likelihood Classification of Quadrature Amplitude Modulation”, *Proc. WCNC*, Sydney, 18-21 Apr. 2010, p. 1-5.

[205] Q. Shi and Y. Karasawa, “Automatic Modulation Identification Based on the Probability Density Function of Signal Phase”, *IEEE Transactions on Communications*, Vol. 60, No. 4, Apr. 2012, p. 1033-1044.

[206] C.-Y. Huang and A. Polydoros, “Likelihood Methods for MPSK Modulation Classification”, *IEEE Transactions on Communications*, Vol. 43, No. 2/3/4, Feb/Mar./Apr. 1995, p. 1493-1504

[207] C. Long, K. Chugg and A. Polydoros, “Further Results in Likelihood Classification of QAM Signals”, *Proc. MILCOM*, Fort Monmouth, 2-5 Oct. 1994, p. 57-61.

[208] B. Dulek, O. Ozdemir, P. K. Varshney and W.Su, “A Novel Approach to Dictionary Construction for Automatic Modulation Classification”, *Journal of the Franklin Institute*, Vol. 351, 2014, p. 2991-3012.

[209] M. L. D. Wong and A. K. Nandi, “Semi-Blind Algorithms For Automatic Classification of Digital Modulation Schemes”, *Digital Signal Processing*, Vol. 18, 2008, p. 209-227.

[210] M. S. Muhlhaus, M. Oner, O. A. Dobre and Friedrich K. Jondral, “A Low Complexity Modulation Classification Algorithm for MIMO Systems”, *IEEE Communications Letters*, Vol. 17, No. 10, Oct. 2013, p. 1881-1884.

[211] L. Tong and S. Perreau, “Multichannel Blind Identification: From Subspace to Maximum Likelihood Methods”, *Proceedings of the IEEE*, Vol. 86, No.10, 1998, p.1951-1968.

[212] M. Abuthinien, S. Chen and L. Hanzo, “Semi-Blind Joint Maximum Likelihood Channel Estimation and Data Detection for MIMO Systems”, *IEEE Signal Processing Letters*, Vol. 15, 2008, p. 202-205.

[213] O. Bayer and M. Oner, “Joint Space Time Block Code and Modulation Classification for MIMO Systems”, *IEEE Wireless Communications Letters*, Vol. 6, No. 1, Feb. 2017, p. 62-65.

[214] R. Prasad, C. R. Murthy and B. D. Rao, “Joint Channel Estimation and Data Detection in MIMO-OFDM Systems: A Sparse Bayesian Learning Approach”, *IEEE Trans. on Signal Processing*, Vol. 63, No. 20, Oct. 2015, p. 5369-5382.

[215] A. Gomaa, L. M. A. Jalloul, M. M. Mansour, K. Gomadam, and D. Tujkovic, “Max-Log-MAP Optimal MU-MIMO Receiver for Joint Data Detection and Interferer Modulation Classification”, *IEEE Communications Letters*, Vol. 20, No. 7, Jul. 2016, p. 1389-1392.

[216] M. Turan, M. Oner and H. A. Cirpan, “Joint Modulation Classification and Antenna Number Detection for MIMO Systems”, *IEEE Communications Letters*, Vol. 20, No. 1, Jan. 2016, p. 193-196.

[217] C. Cozzo and B. L. Hughes, “An Adaptive Receiver for Space-Time Trellis Codes Based on Per-Survivor Processing”, *IEEE Transactions on Communications*, Vol. 50, No. 8, Aug. 2002, p. 1213-1216.

[218] N. Seshadri, “Joint Data and Channel Estimation and Using Blind Trellis Search Techniques”, *IEEE Transactions on Communications*, Vol. 42, No. 2/3/4, Feb/Mar/Apr 1994, p. 1000-1011.

[219] S. Chen and Y. Wu, “Maximum Likelihood Joint Channel and Data Estimation Using Genetic Algorithms”, *IEEE Transactions on Signal Processing*, Vol. 46, No. 5, May 1998, p. 1469, 1473.

[220] M. J. Omidi, P. G. Gulak and S. Pasupathy, “Parallel Structures for Joint Channel Estimation and Data Detection over Fading Channels”, *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 9, Dec. p. 1616-1629.

[221] E. Soltanmohammadi and M. Naraghi-Pour, “Blind Modulation Classification over Fading Channels using Expectation-Maximization”, *IEEE Communications Letters*, Vol. 17, No. 9, Sep. 2013, pp. 1692-1695.

[222] B. Dulek, “Online Hybrid Likelihood Based Modulation Classification Using Multiple Sensors”, *IEEE Transactions on Wireless Communications*, Vol. 16, No. 8, Aug. 2017, p. 4984-5000.

[223] O. Ozdemir, T. Wimalajeewa, B. Dulek, P. K. Varshney and W. Su, “Asynchronous Linear Modulation Classification With Multiple Sensors via Generalized EM Algorithm”, *IEEE Transactions on Wireless Communications*, Vol. 14, No. 11, Nov. 2015, p. 6389-6400.

[224] B. Shim and I. Kang, “Joint Modulation Classification and Detection Using Sphere Decoding”, *IEEE Signal Processing Letters*, Vol. 16, No. 9, Sep. 2009, p. 778-781.

[225] M. M. Mansour, “Near-ML MIMO Subspace Detection Algorithm”, *IEEE Signal Processing Letters*, Vol. 22, No. 4, Apr. 2015, p. 408-412.

[226] H. Sarieddeen, M. M. Mansour and A. Chehab, “Modulation Classification via Subspace Detection in MIMO Systems”, *IEEE Communications Letters*, Vol. 21, No. 1, Jan. 2017, p. 64-67.

[227] K. Kumar, A. Prakash and R. Tripathi, “Spectrum Handoff in Cognitive Radio Networks: A Classification and Comprehensive Survey”, *Journal of Network and Computer Applications*, Vol. 61, 2016, p. 161-188.

[228] D.-J. Lee and W.-Y. Yeo, “Channel Availability Analysis of Spectrum Handoff in Cognitive Radio Networks”, *IEEE Communications Letters*, Vol. 19, No. 3, Mar. 2015, p. 435-438.

[229] C. Cordeiro, K. Challapali and D. Birru, “IEEE 802.22: An Introduction to the First Wireless Standard based on Cognitive Radios,” *Journal of Communications*, Vol. 1, No. 1, Apr. 2006, p. 38-47.

[230] H.-H. Choi, K. Jang and Y. Cheong, “Adaptive Sensing Threshold Control Based on Transmission Power in Cognitive Radio Systems”, *Proc. CrownCom*, 15-17 May 2008, Singapore, p. 1-6.

[231] N. I. Miridakis, T. A. Tsiftsis, G. C. Alexandropoulos and M. Debbah, “Simultaneous Spectrum Sensing and Data Reception for Cognitive Spatial Multiplexing Distributed Systems”, *IEEE Transactions on Wireless Communications*, Vol. 16, No. 5, May 2017, p. 3313-3327.

[232] J. F.-Segura and X. Wang, “GLRT-Based Spectrum Sensing for Cognitive Radio with Prior Information”, *IEEE Transactions on Communications*, Vol. 58, No. 7, Jul. 2010, p. 2137-2146.

[233] R. Zhang, T. J. Lim, Y.-C. Liang and Y. Zeng, “Multi-Antenna Based Spectrum Sensing for Cognitive Radios: A GLRT Approach”, *IEEE Transactions on Communications*, Vol. 58, No. 1, Jan. 2010, p. 84-88.

[234] P. Wang, J. Fang, N. Han and H. Li, “Multiantenna-Assisted Spectrum Sensing for Cognitive Radio”, *IEEE Transactions on Vehicular Technology*, Vol. 59, No. 4, May, 2010, p. 1791-1880.

[235] M. H. Hayes, “*Statistical Digital Signal Processing and Modeling*”, John Wiley & Sons, 1996.

[236] B. F.-Boroujeny, “Filter Bank Spectrum Sensing for Cognitive Radios”, *IEEE Transactions on Signal Processing*, Vol. 56, No. 5, May 2008, p. 1801-1811.

[237] S. Haykin, D. J. Thomson and J. H. Reed, “Spectrum Sensing for Cognitive Radio”, *Proceedings of the IEEE*, Vol. 97, No. 5, May 2009, p. 849-877.

[238] J. Wang and Q. T. Zhang, “A Multitaper Spectrum Based Detector for Cognitive Radio”, *Proc. WCNC*, 5-8 Apr. 2009, Budapest, p. 1-5.

[239] Q. Zhang, “Multitaper Based Spectrum Sensing for Cognitive Radio: Design and Performance”, *Proc. VTC*, 15-18 May 2011, Yokohama, p. 1-5.

[240] E. H. Gismalla and E. Alsusa, “New and Accurate Results on the Performance of the Multitaper-Based Detector”, *Proc. ICC*, 10-15 Jun. 2012, Ottawa, p. 1609 – 1613.

[241] E. H. G. Yousif, T. Ratnarajah and M. Sellathurai, “Modeling and Performance Analysis of Multitaper Detection Using Phase-Type Distributions Over MIMO Fading Channels”, *IEEE Transactions on Signal Processing*, Vol. 63, No. 22, Nov. 2015, p. 5882-5896.

[242] Q. T. Zhang, “Theoretical Performance and Thresholds of the Multitaper Method for Spectrum Sensing”, *IEEE Transactions on Vehicular Technology*, Vol. 60, No. 5, Jun. 2011, p. 2128-2138.

[243] V. Kuppusamy and R. Mahapatra, “Primary user detection in OFDM based MIMO Cognitive Radio”, *Proc. CrownCom*, 15-18 May 2008, Singapore, p. 1-5.

[244] A. Nafkha and B. Aziz, “Closed-Form Approximation for the Performance of Finite Sample-Based Energy Detection Using Correlated Receiving Antennas”, *IEEE Wireless Communications Letters*, Vol. 3, No. 6, Dec. 2014, p. 577-580.

[245] D. R. Joshi, D. C. Popescu and O. A. Dobre, “Dynamic Threshold Adaptation for Spectrum Sensing in Cognitive Radio Systems”, *Proc. RW*, 10-14 Jan. 2010, New Orleans, p. 468-471.

[246] N. Wang and Y. Gao, “Optimal Threshold of Welch’s Periodogram for Sensing OFDM Signals at Low SNR Levels”, *Proc. EW*, 16-18 Apr. 2013, Guildford, p. 1-5

[247] M. Namdar, H. Ilhan and L. D.-Ata, “Dispersed Chirp-z Transform-based Spectrum Sensing and Utilisation in Cognitive Radio Networks”, *IET Signal Processing*, Vol. 8, Iss. 4, Jun. 2014, p. 320-329.

[248] M. Namdar, H. Ilhan and L. D.-Ata, “Optimal Detection Thresholds in Spectrum Sensing with Receiver Diversity”, *Wireless Personal Communications*, Vol. 87, No. 1, 2016, p. 63-81.

[249] N. Wang, Y. Gao and X. Zhang, “Adaptive Spectrum Sensing Algorithm Under Different Primary User Utilizations”, *IEEE Communications Letters*, Vol. 17, No. 9, Sep. 2013, p. 1838-1841.

[250] S. Zhang and Z. Bao, “An Adaptive Spectrum Sensing Algorithm under Noise Uncertainty”, *Proc. ICC*, 5-9 Jun. 2011, Kyoto, p. 1-5.

[251] S. Bokharaiee, H. H. Nguyen and E. Shwedyk, “Blind Spectrum Sensing for OFDM-Based Cognitive Radio Systems”, *IEEE Transactions on Vehicular Technology*, Vol. 60, No. 3, Mar. 2011, p. 858-871.

[252] Z. Quan, S. Cui, A. H. Sayed and H. V. Poor, “Optimal Multiband Joint Detection for Spectrum Sensing in Cognitive Radio Networks”, *IEEE Transactions on Signal Processing*, Vol. 57, No. 3, Mar. 2009, p. 1128-1140.

[253] T. E. Bogale, L. Vandendorpe and L. B. Le, “Wide-Band Sensing and Optimization for Cognitive Radio Networks with Noise Variance Uncertainty”, *IEEE Transactions on Communications*, Vol. 63, No. 4, Apr. 2015, p. 1091-1105.

[254] Y.-C. Liang, Y. Zeng, E. C.Y. Peh and A. T. Hoang, “Sensing-Throughput Tradeoff for Cognitive Radio Networks”, *IEEE Transactions on Wireless Communications*, Vol. 7, No. 4, Apr. 2008, p. 1326-1337.

[255] M. Guerrini, L. Rugini and P. Banelli, “Sensing-Throughput Tradeoff for Cognitive Radios”, *Proc. SPAWC*, 16-19 Jun. 2013, Darmstadt, p. 115-119.

[256] Z. Bai, L. Wangand and K. S. Kwak, “Different Sensing Durations-Based Cooperative Spectrum Sensing in Cognitive Radio Systems”, *Wireless Communications and Mobile Computing*, Vol. 14, 2014, p. 1522–1529.

[257] T.-W. Chiang, J.-M. Lin and H.-P. Ma, “Optimal Detector for Multitaper Spectrum Estimator in Cognitive Radios”, *Proc. GlobeCom*, Honolulu, 30 Nov. - 4 Dec. 2009, p. 1-5

[258] M. K. Jataprolu, R. D. Koilpillai and S. Bhashyam, “Optimal MTM Spectral Estimation Based Detection for Cognitive Radio in HDTV”, *Proc. NCC*, Kharagpur, , 3-5 Feb. 2012, p. 1-5.

[259] J. G. Proakis and D. G. Manolakis, “*Digital Signal Processing Principles, Algorithms, and Applications*”, Prentice Hall, 1996.

Appendix (A)

The Gauss-Hermite quadrature is a form that can be used to approximate the problem of (5.10), check [139] and the online sources below for further details.

- http://www.efunda.com/math/num_integration/findgausshermite.cfm
- http://en.wikipedia.org/wiki/Gauss%E2%80%93Hermite_quadrature

This quadrature form is used to approximate complex integrals of the following shape

$$\int_{-\infty}^{\infty} e^{-z^2} f(z)dz = \sum_{i=1}^{n} w_i f(z_i) \tag{A.1}$$

where n is the number of samples used in the approximation. The z_i are the roots of the Hermite polynomial $H_i(z)(i = 1,2,\dots,n)$, and the associated weights w_i are given by

$$w_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 [H_{n-1}(z_i)]^2} \tag{A.2}$$

The approximation parameters for 16 points used in Chapter Five, are tabulated below

Table 5.4 Gauss-Hermite quadrature approximation using 16 points.

No	z_i abscissas	w_i weights
1	-4.688738939	2.65^{-10}
2	-3.869447905	2.32^{-07}
3	-3.176999162	2.71^{-05}
4	-2.546202158	0.000932284
5	-1.951787991	0.012880312
6	-1.380258539	0.083810041
7	-0.822951449	0.280647459
8	-0.273481046	0.507929479
9	0.273481046	0.507929479
10	0.822951449	0.280647459
11	1.380258539	0.083810041
12	1.951787991	0.012880312
13	2.546202158	0.000932284
14	3.176999162	2.71^{-05}
15	3.869447905	2.32^{-07}
16	4.688738939	2.65^{-10}

Appendix (B)

Pseudocode for the MTSE-STBC and PSE-STBC Algorithms.

1: **Initialise** constants (M=2 for BPSK, K=1024, L=3) and Monte Carlo = 10^5

2: **Set** α=0.5, P_{md}=0.01, and P_{fa}=0.01→ $\acute{\eta}$, $N_r \times N_t$=2 × 2, 4 × 4, 8 × 8

3: **Generate** the BPSK Sequence

4: **For** 1: N_r (Receive Antenna)

5: **For** SNR [-30 to 30] dB

6: **For** Repetition [1 to Monte Carlo]

7: **Generate** AWGN of σ_n^2=1

8: **Generate** Flat-Fading Rayleigh Channel of σ_h^2=1 Equation (3.4)

9: **For** 1: N_t (Transmit Antenna)

10: Compute STBC Signal with AWGN and Rayleigh Channel Equation (7.1), or (3.3) or (3.21)

11: Equalise Received STBC Signal using Channel Conjugate and MRC Equation (3.22) or (3.24)

12: **End For**

13: **Compute** MTSE-STBC Equation (7.22)

14: **Compute** PSE-STBC Equation (7.22) without Tapers

15: **Apply** BHT per Equation (7.27)

16: **Calculate** Error Sensing Probability

17: **End For**

18: **End For**

19: **End For**

20: **Display** Results

Author's List of Publications

A. Journals

1. Ahmed O. Abdul Salam, Ray E. Sheriff, Yim F. Hu, Saleh R. Al-Araji and Kahtan Mezher, "Adaptive IMM-Kalman Filter for Efficient Modulation Classification in Cognitive Radios", *Submitted to IEEE Transactions on Vehicular Technology.*
2. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, "Spectrum Estimation in Cognitive Radio Based on Multitaper and Higher-Order STBC Techniques and Featuring Threshold Adjustment", *Submitted to IEEE Transactions on Mobile Computing.*
3. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, "Spectrum Estimation in Cognitive Radio Based on MIMO-OFDM and Multi-taper Techniques", *Submitted to Annals of Telecommunications.*
4. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, "Adaptive Threshold and Optimal Frame Duration for Multi-taper Spectrum Sensing in Cognitive Radio", *ICT Express, in press*.
5. A. O. Salam, R. E. Sheriff, S. R. Al-Araji, K. Mezher and Q. Naisr, "Adaptive Interacting Multiple Model-Kalman Filter for Multitaper Spectrum Sensing in Cognitive Radio", *Electronics Letters*, Vol. 54, Iss. 5, Mar. 2018, pp. 321-322.
6. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, "A Multi-Taper Spectrum Based Estimator for Cognitive Radio Using Multiple Antennas and Space-Time Block Code Techniques", *IET Circuits, Devices and Systems*, Vol. 12, No. 2, Mar. 2018, pp. 133-143.
7. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, "Novel Approach for Modeling Wireless Fading Channels Using a Finite State Markov Chain", *ETRI*, Vol. 39, No. 5, Oct. 2017, pp. 718-728

B. Conferences

1. Ahmed O. Abdul Salam, Ray E. Sheriff, Yim F. Hu, Saleh R. Al-Araji and Kahtan Mezher, “Assessment on Using Multitaper and Higher-Order STBC Techniques for Spectrum Estimation in Cognitive Radio”, *IEEE VTC-Fall*, Toronto, Sep. 2017.
2. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, “Multi-Taper and MIMO Techniques for Spectrum Sensing in Cognitive Radio”, *IEEE ICECS*, Cairo, Dec. 2015.
3. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, “Automatic Modulation Classification in Cognitive Radio Using Multiple Antennas and Maximum-Likelihood Techniques”, *IEEE CIT*, Liverpool, Oct. 2015.
4. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, “A Unified Practical Approach to Modulation Classification in Cognitive Radio using Likelihood-Based Techniques”, *IEEE CCECE*, Halifax, Jun. 2015.
5. Ahmed O. Abdul Salam, Ray E. Sheriff, Saleh R. Al-Araji, Kahtan Mezher and Qassim Naisr, “An Overview on Non-Parametric Spectrum Sensing in Cognitive Radio”, *IEEE ICCES*, Cairo, Dec. 2014.
6. Ahmed O. Abdul Salam, Saleh R. Al-Araji, Kahtan Mezher, Qassim Naisr and Ray E. Sheriff “A General Perspective on Software-Hardware Defined Cognitive Radio Based on Emergency Ad-Hoc Network Topology”, *IEEE IHTC*, Montreal, Jun. 2014.